

nature**29 November 1979**

Universities are an international heritage

OVERSEAS students in British universities, whether at the undergraduate or postgraduate level, will have to pay more — much more — for their education from next September. The fees will vary from subject to subject but will range from £2,000 to £5,000 per year; at the same time, sources of assistance for students in developing countries, whose governments are unlikely to be able to find the additional sums, are drying up with the decline in scholarships that the British Council can provide on its trimmed budget.

Not surprisingly, in the weeks since the announcement was made arguments about the folly of the increase have been widely aired — particularly those which point to the financial benefits which accrue to Britain from overseas students. Against these are ranged arguments that if British education is all that good, market-place forces will continue to bring overseas students in droves; and if, on the other hand, it comes off badly in international comparisons, then again the market-place will convey a clear message and we will all have to pull our socks up.

The trouble with the market-place argument is that whilst the prospective purchasers of everything from toothpaste to automobiles can at leisure compare costs and assess values, the specific question of which university in which country is likely to provide the best (whatever that means) education for an

individual student is almost unanswerable. Decisions are based as much on tradition and hearsay as anything else. Maybe for this reason, good students will continue to be sent to Britain. No-one yet knows. But whatever happens it is foolish to talk of higher education existing in anything like a market-place.

The real concern over students' fees ought to be the implications they convey, not altogether unjustly, of a developing defensiveness in Britain against foreigners taking any advantage whatsoever of the British taxpayer. Traditionally, British governments have recognised the intangible benefits which come from a generous policy towards students — not least that places of learning are an international, not just a national, heritage and that a leavening of overseas students helps keep them that way. If the present government is seen simply to be milking overseas students for what it can get, then, no matter whether enough rich students can be found from overseas to keep the numbers up, the character of Britain's universities will slowly change for the worse.

Monitoring developments in numbers of applicants is not enough. It says nothing about those who have been frightened off by the price from filling in a form. The government needs to take remedial action now, probably in the form of a scholarship fund, to ensure that deserving students from poorer countries get a chance. □

Threshold treaty would be a mess

PROGRESS these past few months towards a ban on underground nuclear weapons tests has been glacial. The United States, Soviet Union and Britain continue to meet regularly in Geneva to discuss technical matters such as the number of monitoring stations that each country will have, what instruments the stations will contain and so on. But there seems little sense of urgency, despite the obloquy that nuclear powers are bound to encounter during the Non-Proliferation Treaty Review Conference next year if they have failed to deliver on a treaty which has a high symbolic value to non-nuclear states.

The idea that the treaty might be of indefinite duration has long since been abandoned; a short-term ban of three years with no promise whatsoever of renewal is now the most that will ever be negotiated. But even that now shows danger of being eroded. There are rumours that whilst Britain is at present maintaining an official line that a total ban is still the desirable outcome, privately views are being expressed that a much better treaty would be one which kept the concept of a threshold. A figure mentioned is 15 kilotons. Explosions with yields above this would be forbidden,

explosions with lower yields would not.

Technically, this sort of treaty would be next to impossible to police properly. At present the United States and Soviet Union have a threshold agreement that neither will fire devices of more than 150 kilotons, and even at this level there have been occasions when serious questions have been asked about whether particular tests have contravened the limit: such are the vagaries of geology. A reduction of this limit by a factor of ten would cause much greater difficulty, because signal levels recorded by various seismic stations around the world would then be closer to the level of the general noise background. As a consequence the error bars that have to be assigned to yield estimates would expand unacceptably; it would be a foolhardy person who was prepared to say that any particular device was, say, 20 kilotons rather than 15. The consequences of so many false alarms are not difficult to imagine.

There are strong political reasons why a comprehensive test ban should be comprehensive. There are equally strong technical ones. □



US debate heats up over energy conservation strategies

Higher energy prices may encourage more efficient use — but will the poor have to bear an unfair proportion of the costs? **David Dickson** reports

THE recent take-over of the US Embassy in Teheran has focused Washington's attention even more sharply on the need to reduce dependence on imported oil. And this in turn has emphasised the gains to be made through measures to increase energy conservation.

Yet the appropriate strategy to achieve energy savings has become a major source of political controversy. At stake is the basic trade-off between energy efficiency and energy equity: if the price of energy is raised to encourage conservation, can this be done without imposing an unfair burden on those who would be proportionately most affected by the higher costs, namely the urban and rural poor for whom energy costs take up to 30 to 40% of the weekly income?

This conflict has been most visible in early sparring between the two main contenders for next year's Democratic presidential nomination. President Carter has made oil price de-control central to his energy strategy, arguing that the price of domestic oil should be allowed to rise to that set by the OPEC countries — and that there should be a 'windfall' profits tax on the oil companies from which the worst effects of the price increase can be mitigated.

In contrast, Senator Edward Kennedy has in recent weeks been arguing strongly against decontrol, suggesting that in current circumstances 'it would seriously exacerbate our inflation and produce 'conservation' only as a by-product of a massive recession'. Higher prices may in the long run be desirable to encourage economic competition — but in the present economic climate they would be unjust.

Not that the Carter administration has been idle on conservation. Last week, for example, the Department of Energy published proposed building regulations estimated to add 3 to 5% to the costs of new residential and commercial buildings — but to result in energy savings of 20 to 50% through setting overall consumption rates needed to qualify for federal loan guarantees and other benefits.

But after setbacks in other parts of its proposed policy, the administration has recently switched its main focus to energy supply matters, such as the proposed massive programme to develop synthetic fuels. And the conservation ball has been



The American poor have little money left for heat or light

passed back to Congress, where the Senate's Labour and Human Resources Committee is supporting a bill that would provide \$2.2 billion within three years to help poor people insulate their homes.

This see-sawing between the administration and Congress reflects the difficulties that surround attempts to devise a strategy for conservation, where the unpredictable effects of policy measures — as well as the complexity of putting them into practice — are even greater than for energy supply decisions.

Indeed it is only relatively recently that the full potential of conservation for reducing energy demand, not only for transportation but also for space heating and other domestic uses, has been accepted in policy circles. And there has been a tendency to lean heavily on the effects of 'price signals' — the expected response of individuals to increased energy costs.

Using such signals, for example, a study prepared two years ago for the National Academy of Sciences' Committee on Nuclear and Alternative Energy Systems (CONAES), produced estimates showing the wide range of energy-use forecasts resulting from different pricing strategies. It predicted, for example, that a four-fold increase in energy costs would lead to energy use remaining approximately constant between 1975 and 2010.

In practice, however, price signals have so far not provided a particularly good guide to the behaviour of car drivers, either in seeking energy-efficient vehicles or in cutting down consumption of petrol. "I am probably less optimistic now about how readily people take to economic signals", Dr John Gibbons, chairman of the NAS panel and now director of the Office of Technology Assessment, said last week.

"I am still convinced that people make rational economic decisions about future choices — but it seems that you still have a

lot of social inertia, perhaps because the path to high energy efficiency is complex".

Experience has therefore shown that the relationship between price and consumption is even more elastic than previously thought. This has led to greater consideration of non-price mechanisms, inevitably involving greater federal intervention. At the same time energy problems have become intertwined with broader attempts to cope with inflation.

Here the choice is whether, as Senator Kennedy argues, the government should tackle inflation head-on by trying to keep prices down until the economic situation improves — or to argue that eventually inflation can only be contained by realistic pricing strategies, and that the government's role should be to stimulate the operation of market principles.

"In the long run, it is going to hurt if prices stay down", says Dr Lee Shiffer of the Lawrence Berkeley Laboratory in California, a member of the panel which produced the conservation report for CONAES. "One thing that our study found was that for the majority of consumers, energy standards will only be accepted if . . . the savings are in the same ball-park as the costs."

The panel's calculations show that only when the cost of oil rises to about \$30 a barrel do major conservation programmes begin to look attractive. And this strategy is also supported by many environmentalists, who have argued that increased conservation will not only slow the growth of pollution-generating sources, but will also spur the search for more desirable sources of energy, such as solar power.

Such arguments, however, are meeting increasing opposition from community activists and others concerned more directly with social welfare issues. Gar Alperovitz, for example, of the Center for Economic Alternatives in Washington,

argues that whatever the theoretical principles involved, steps to offset the inflationary impact of increased energy prices are unlikely to prevent the poor from being required to shoulder a disproportionate amount of the burden.

The effects of recent price increases on such groups is illustrated by a survey prepared for the office of consumer affairs of the Department of Energy. This points out that between 1973 and 1977, the median income of poverty-level families increased 34.7%, while the costs of household fuels increased 78%, and has risen even more dramatically since then.

The survey found that for many poor

families, food was the first budget item to be cut in order to pay high utility bills. "Since few of them spend much money on anything except housing and food, there is nothing else to eliminate . . . many of them are without heat at all in the winter", the survey says.

If nothing else, current debates have raised a general awareness that energy politics have important equity implications, and cannot be resolved merely by technical arguments. But the tensions between environmentalists and community activists is symptomatic of the disagreement that remains over appropriate strategy.

Many are now demanding a stronger government role in regulating demand, through rationing, heavy energy taxes, and tough conservation standards. They point, for example, to the example of Portland, Oregon, which has recently introduced a statute which requires all home owners to introduce conservation measures if they want to sell or to rent their properties.

"We cannot realistically overcome the inequity of higher energy prices" says one community worker from New York. But with Congress in its current anti-regulatory mood, alternative strategies are unlikely to get very far, and protection for the poor will have to be sought by other routes. □

Canada boosts university research by 32%

THE Canadian government has announced that the budget of the Natural Sciences and Engineering Research Council (NSERC) will be increased by 32% in the financial year 1980-81, to a total of \$159.8 million.

The increase is the first of a number of measures expected to be made public in the next few weeks for stimulating Canadian research and development. This is in line with the government's pre-election commitment, reaffirmed two weeks ago by the newly-appointed Minister for Science and Technology, Mr Heward Grafftey, to raise Canadian R & D expenditure from 1% to 2.5% of the gross national product.

The extra money has been provided to the research council, which is responsible for supporting research in Canadian universities, in response to a five-year plan which the council submitted to the government earlier this year.

"We are currently operating with a budget which is 14% lower in real terms than it was ten years ago. This is obviously a very significant increase, although still less than we had asked for", Dr George Macnabb, President of NSERC, told *Nature* last week.

In response to current and projected manpower shortages, the council is planning to introduce a research associate scheme, under which it will provide salary costs and additional research support over a five-year period to post-doctoral scientists at universities prepared to commit 90% of their time to research. About 100 such awards will be made annually.

Another plan is to provide summer support for undergraduates in their second or third year who want to gain research experience. The council is currently working out details on how the extra money for equipment should be distributed. "We are trying to undo the damage of the last decade, and hopefully we can do it in a planned way", says Dr Macnabb.

As for research grants, the precise distribution of awards will not be announced until February, but it is planned that although the major part of the increase will go to basic research, there will be propor-

tionately a larger increase in support for targeted research.

The NSERC'S budget increase has been generally welcomed by Canadian scientists, who in recent months have been lobbying for an end to the decline in federal support for research through groups such as the Canadian Federation of Biological Societies and the Canadian Association of Physicists (CAP).

"The federal commitment to improve support for fundamental research in the universities and to ensure the creation of a healthy and vital university research community is long overdue", Dr Robert Willes of the CFBS commented last week.

Dr John Kucharczyk of CAP, however, warned that scientists still had some reservations. These included: doubts

whether the commitment to increased funding would continue for more than one year; doubts whether the provinces and private industry were prepared to provide the support on which the government's strategy depended for research; and concern over the constraints placed on the NSERC by the fact that less than half its requested budget increase had been accepted.

The Canadian government is expected to announce shortly increased funding for both the Medical Research Council and the Social Science Research Council. And the budget to be presented to Parliament in December is likely to include details of tax incentives offered to industry to invest in research and development.

David Dickson

US drops back in biology and maths

THE total number of US papers appearing in the world's leading scientific publications fell by 6% between 1973 and 1977, according to figures released in Washington last week by the National Science Board, the policy arm of the National Science Foundation.

US prominence dropped particularly in biology and mathematics. In biology, the net number of papers by US scientists in the Science Citation Index's two thousand most influential journals fell by 11%, representing a drop from 46 to 42% of the total biology papers. In mathematics, the proportion of US-authored papers fell from 48 to 41% over the four years.

This decline, which was mirrored but not quite so sharply in almost all scientific disciplines, was partly offset by an increase in the average frequency with which papers published were cited. In 1973 the rate of citation of US papers was 28% above what would have been expected on purely numerical grounds. By 1977 this had risen to 30%.

These figures, based on data gathered by Computer Horizons Inc. from the Institute for Scientific Information, are published in *Science Indicators 1978*, presented to Congress last week by President Carter as

the annual report of the National Science Board.

Despite frequent statements to the contrary, the report paints a relatively healthy picture of US research and development, suggesting that an apparent decline in international terms may be largely the result of countries such as Japan and West Germany catching up with the US's previously unchallenged position.

"Among the two major types of research activity — basic and applied — a period of modest growth has persisted for longer than any time since the late 1960s" the report says. Even in industry, it states, expenditure on basic research grew from \$719 million in 1975 to \$975 million in 1978, an increase in constant dollars of 14% over the years.

The report points out that the "phenomenal growth" which has taken place in energy research and development since 1974 — an occurrence which, contrary to expectations, it says did not take place in other industrialised countries — has slowed down, although perhaps only temporarily. But a shift from military to civilian R & D within the federal budget seem to have been reversed in the last two years. □

Defoliant companies can be sued

A US federal judge ruled last week that five chemical companies which produced the defoliant "agent orange" can be sued by Vietnam veterans. The veterans claim that they were injured by exposure to the chemical, and that their children were deformed. Agent Orange contains the defoliant 2,4,5-T, suspected of containing the toxic contaminant dioxin; it was extensively used in Vietnam between 1962 and 1971.

The veterans want a trust fund to be set up and financed by the five chemical companies — Dow Chemical, Monsanto, Thompson Hayward Chemical Company, Hercules Inc, and Diamond Shamrock — for the benefit of those who, they claim, suffered as a direct or indirect result of exposure to the defoliant. According to the veteran's attorney, claims against the companies could be \$40,000 million.

The companies had contended that the 362 veterans are entitled only to veteran's medical benefits, and that the suits should be tried individually in state courts. However, Judge George C Pratt ruled that a

suit could be brought against the companies.

The US Defense Department was accused last week of severely underestimating the number of soldiers exposed to Agent Orange during operations in Vietnam. Although the Pentagon has until recently said that all combat zones sprayed with the defoliant were free of army personnel, a report issued last Saturday by the General Accounting Office, the investigatory arm of Congress, said that at least 5,900 marines were within a third of a mile of the areas while or immediately after they were being sprayed. The GAO also says that a further 16,000 marines were within a third of a mile of the sprayed areas within four weeks of the spraying, and it called the army's contention that troops did not enter the area until six weeks afterwards as "inaccurate". In a letter to Senator Charles Percy who had requested the GAO investigation, the Department of Defense admitted that "the chances that ground troops were exposed to herbicide Orange are higher than DOD previously acknowledged". □

UK radical scientists ten years on

THE 10th anniversary of the founding of the British Society for Social Responsibility in Science was celebrated in London last week by a congenial gathering of 200 activists representing a span of 60 years of experience of radical and socialist approaches to science. Five speakers analysed the past and present successes of the "largest single grouping of radical scientists in Europe".

Maurice Wilkins, president of the society, evoked the social atmosphere requiring new approaches to the practice of science by emphasising the turbulent state of world affairs. "Things have changed a very great deal in a very short time. We have seen a continuous disintegration and degeneration of our institutions in modern society over the past 50 years. This disintegration has certain advantages since it increases the possibility of getting outside our conditions. It is difficult to overestimate how strong our social conditioning is and BSSRS correctly has been concerned to show that science cannot escape from these conditions."

Joseph Needham, who is as close to a patron saint as is possible for a political organisation to have, presented a unique mixture of Christianity and socialism with a blessing for BSSRS' past and future work. He affirmed his conviction in the achievements of socialism. "Socialism has been an outstanding success in China which was absolutely medieval in character in the 1940s. I think we have to recognise that India will never get anywhere unless it follows the Chinese way."

Strong continuities exist between the issues of the 30s science movement and

those of present groups. Racism, science in war, organising scientific workers, developing links with left wing movements, demystifying science to a larger audience and the position of women in science are all issues that lie outside mainstream "establishment science". One of the great achievements of the 30s movement according to Needham was the campaign leading to the eventual construction of the Anderson air-raid shelters prior to the second world war. "We knew it was important and possible and in spite of government indifference we shamed them into doing it."

Speakers, Dorothy Griffiths, Jonathan Rosenhead and Mike Cooley outlined BSSRS recent history and immediate future plans. In terms of analysis, the group has moved away from the "use-abuse" model of science and technology and is concentrating on developing its understanding of how science is thoroughly integrated into the social fabric through its employment in profitable economic production, its use in social control technology and its use as "a justification for social inequality based on class, race and sex." New activities include a formal Women's Caucus, a working group to criticise the UK government's proposal to count and predict the size of Britain's black population, particularly the proposed introduction of a question on racial origin in the 1981 census, and a working group that will provide an information service to trade unions on the likely impact of the proposed introduction of microprocessor based technologies in particular workplaces. **Joe Schwartz**

London meeting calls for end to psychiatric abuse

THE recent meeting in London of the World Psychiatric Association focused attention, once again, on the use of psychiatry for political repression. Although the 1977 Honolulu Congress of the WPA formally condemned psychiatric abuse, there has so far been little concrete action from its special "Committee to Review the Abuse of Psychiatry". *Ad hoc* pickets at the meeting and Amnesty International in its formal letter to the WPA executive, stressed the need for firm and expeditious action from the Review Committee.

The Amnesty appeal stressed the "consistent pattern" of psychiatric abuse in Romania and the Soviet Union, particularly the latter. Details of over 40 recent Soviet cases were appended, with "symptoms" ranging from Seventh Day Adventism (Eugenius Darguzes) to membership of an unofficial Marxist discussion group (Aleksandr Skobov) and from organising an independent trade union (Vladimir Klebanov) to proposing Academician Andrei Sakharov as a candidate for the election to the Supreme Soviet (Vladimir Tsurikov). In spite of the recent criticism of "Snezhnevskiiism" (*Nature* 11 October, page 419) Snezhnevskii-type diagnoses of "sluggish schizophrenia" continue — as does the "treatment" of dissent by massive doses of neuroleptic drugs.

Perhaps one of the most telling comments on psychiatric abuse is the fact that the occasional transfer of a dissident such as M Kukobaka (charged with "anti-Soviet slander") from psychiatric hospital to trial and prison is viewed by Soviet human rights campaigners as a qualified triumph. **Vera Rich**

Galileo formally reprieved

POPE John Paul II has called for a formal reversal of the Catholic Church's verdict on Galileo, who in 1633 was condemned by the Roman inquisition and placed under house arrest for his heliocentric theory of the solar system. The Pope was addressing a special session of the Vatican's Pontifical Academy of Sciences, held to mark the centenary of the birth of Albert Einstein.

He quoted the Pastoral Constitution *Gaudium et Spes* No 36/2 of the second Vatican Council, which deplores "certain habits of mind sometimes found too among Christians, which do not sufficiently attend to the rightful independence of science. The arguments and conflicts which they spark lead many minds to conclude that faith and science are mutually opposed". These words, said the Pope, related in particular to the Galileo case. □

news in brief

US waives air pollution rule for coal-burning power plant: The US Environmental Protection Agency announced last week that it would temporarily waive air pollution rules to allow New England's largest power plant — in Somerset, Massachusetts — to convert from oil to coal two years earlier than planned. The utility had been ordered to convert to coal in 1977 by the Federal Energy Administration, and although the EPA had agreed to the conversion, the need to install pollution-control equipment meant that the conversion was initially planned to be phased in between 1982 and 1984. However the EPA has agreed to a temporary waiver of the pollution rule covering particulate emissions.

Converting most power plants to coal is a key element of the Carter Administration's energy programme, and EPA deputy administrator Barbara Blum said that the waiver, which will not involve the violation of atmospheric air quality standards protecting public health, was an "important milestone" towards energy independence.

British Council of Churches urges abolition of UK nuclear missiles: By a vote of 54-19, the semi-annual Assembly of the British Council of Churches has urged the UK government to abandon its plans to replace Polaris submarine missiles with newer nuclear missile systems. The Assembly resolved that "non-replacement by the UK of its present nuclear strategic deterrent (the Polaris missile system) would strengthen moves for nuclear non-proliferation and urges Her Majesty's government to take a decision to this effect."

In its presentation of the resolution, the Council's Division of International Affairs, argued that "unfavourable consequences" could be minimised by careful implementation of the decision and that there is "no hope of reversing the nuclear arms race until a nuclear weapons state is willing to demonstrate that the arms spiral can go down as well as up. Britain now has a unique opportunity to demonstrate that nuclear non-proliferation is a realistic rather than Utopian goal." At present the UK has four nuclear submarines each armed with 16 missiles which were obtained "at rather advantageous prices" from the US in 1962. Each missile contains three UK manufactured warheads. The US now has 2,141 missiles capable of delivering 11,000 warheads and the Soviet Union has 2,582 missiles with about 5000 warheads.

Jacob criticises new social darwinists: French Nobel Laureate François Jacob has vigorously attacked the "ideological distortions" of biology propagated by social darwinists, racists and others in an attempt to "justify certain models of society". M Jacob is a coauthor of a recent report for the French government on bioengineering and has written a major five part series on biology for the French daily *Le Monde*. In these articles he takes severely to task those who, like the French "new right" or certain sociobiologists, would see in genetic diversity a racial weakness or an excuse to justify inequality and injustice. Indeed, says Jacob, human potential depends on this diversity. "Any effort that aims at the homogenization of the biological properties of individuals — either by hoping to 'ameliorate' them through eugenics or by seeking to encourage one attribute like an aptitude for mathematics or for running — would be biologically suicidal and socially absurd".

Refusnik charged with plagiarism: Dr Vladimir Kislik, a leading figure among Soviet Jewish "refusnik" scientists has been threatened with a charge of plagiarism and breach of copyright. Dr Kislik, whose work has been largely in radiation physics, especially radiation defects in solids, applied for an immigration visa for Israel, in 1973. Permission was refused, on the grounds that radiation physics was "secret", and he was obliged to leave his post at the Kiev Physics Institute. Since then he has worked as a nightwatchman, and as bus-ticket seller, and has been subject to

constant police surveillance. Deprived of any way of continuing his scientific career at home, he managed to send abroad a paper on the kinetics of helium release from irradiated samples of austenite steels, which was published in the *Journal of Nuclear Materials* in 1978.

Recently Dr Kislik was warned that a charge is being prepared against him under article 136 of the Ukrainian penal code which deals with plagiarism and patent infringements. This appears to be in connection with a thermokinetic method of He analysis which he developed while at the Kiev Physics Institute, and with the unauthorised publication of his paper abroad.

This is the first time that the smear of plagiarism has been raised against any refusnik scientist, and appears to form part of a new "pre-Olympic" drive against all forms of dissent and nonconformity. Ironically, Kislik himself could have a strong case against the authorities under the said Article 136. As with Academician Veniamin Levich, after being driven out of academic life, Kislik's name "disappeared" from citation of his published papers.

India's wild life projects

under spy cloud: Many of the wild life projects which India is currently carrying out in collaboration with western countries have come under suspicion of being convenient covers for espionage in the far-flung, strategic borderland of India. The Public Accounts Committee of Parliament, after investigating various collaborative wild life schemes in India, has implied that a project on migratory birds is involved with the perfecting of biological weapons. The US is the guilty partner in this case and the work undertaken in the swampy, sprawling Sunderban, known for its variety of migratory birds, has come under severe public criticism.



"All the beat produced was three KGB, one MIS and a brace of CIA!"

Another project that has become a focal-point of public controversy is the "pheasant project" undertaken in collaboration with the International World Pheasant Association of the UK. It involves the selective reintroduction of pheasant in the foothills of the Himalayas in the states of Uttar Pradesh, Himachal Pradesh and Jammu and Kashmir. It is feared that the foreign collaborators might use the pheasant project as a camouflage for spying in the sensitive borderland.

B Radhakrishna Rao

Four thousand in French protest march: One of the largest demonstrations by scientific and technical researchers in the last ten years marched through Paris on 22 November. Responding to a national call from all the trade unions in the scientific sector to demonstrate opposition to the government's reorganisation plan for the CNRS, scientists and workers from this organisation were joined by colleagues representing the medical, agricultural and overseas research institutions of France in a march from the Sorbonne to the Presidential palace. A delegation handed in a protest against a government reform they feel was pushed through without consultation and which will lead to a dismantlement of French research, overdependency on the criterion of profitability and an increasing level of job insecurity for the country's scientists.

Correction: In "UK to tighten asbestos control", (1 November, page 5) crocidolite should replace chrysotile in the second paragraph of the second column. The first sentence of the fourth paragraph in that column should refer to the 0.5 fibre per ml⁻¹ limit for amosite.

France faces a cold winter

The discovery of cracks in the tubing of most French nuclear reactors has led to a strong call for a national debate on nuclear energy. **Jim Ritter** outlines recent events.

TROUBLES, indeed, come not singly but in armies; and over the past two months France's nuclear programme has found itself under siege by a formidable host of them. The leaks of primary coolant which have, for the moment, shut down half of the country's present generation of nuclear reactors, and the discovery of cracks in the tubing of almost all reactors, either under construction or in operation, promises a cold winter for the French people and a series of agonising reappraisals for their government.

The decision was taken in 1969, under the last government of Pierre Messmer, to create an 'all nuclear' electrical power industry. The development programme was placed in the hands of the state electrical generating board, Electricité de France (EDF) which opted for a two-tier, short-term and long-term programme. The short-term programme involves the building of 900 MW pressurised water reactors under licence from Westinghouse. The long term programme involves the development of fast breeder reactors. The present problems lie in the PWR systems.

By the end of August 1979 the EDF had succeeded in putting into operation six PWR reactors, two at Fessenheim, and four at Bugey near Lyons, the latter alone supplying some 10% of present French electrical generating capacity. Three more PWRs at Gravelines, at Tricastin and at Dampierre were due to be integrated into the network by the end of this year, with another 26 to be added over the next five years (some of these to be of 1300 MW capacity).

But on 22 September the two principal French trade unions, the CFDT Confederation Française Démocratique de Travail (CFDT) and the Confederation Générale de Travail (CGT), sent an open letter to M André Giraud, the Minister for Industry, signalling the existence of numerous cracks in key components of the cooling system of almost all reactors then under construction and, most likely, in those already in service.

The PWR construction firm Framatome had known of these faults since mid-1978 although neither it nor the EDF had made this information public. This the unions meant to remedy. The government's own report appeared on 3 October and was in substantial agreement with that issued by the unions.

Part of the trouble lies in a metal plate, 3.5 m in diameter and 0.53 m thick, used in the steam generation system between the primary and secondary coolants. The plate is pierced by 6776 holes, each 22.22 mm

wide and protected against corrosion by a 10 mm thick layer of Inconel alloy. The alloy is deposited in two stages, both involving the heating of the element to 200°C, generating great thermal stress. It is during the second of these processes, the formation of the Inconel-Inconel bond, that the problem arises. Each plate develops between 30 and 200 cracks from 15 to 20 mm long and 6-8 mm deep. Worse still, it was discovered that similar cracks appeared during the fabrication of the six tubes which connect the reactor chamber with the steam generator, tubes which not only carry the hot, pressurised primary water but also support the entire 400 tonne weight of the reactor itself.

Although these cracks are not dangerous at this stage, Framatome experiments show that they tend to enlarge under the temperature transients of normal operating conditions, leading to possible component failure within 4-6 years, a failure which clearly would have catastrophic consequences involving the total loss of primary coolant.

What then can be done? The steam generating plates are replaceable, though not those in the six reactors already in operation. For these, and for the entry and exit primary tubes, repairs will have to be effected *in situ*. But ambient radiation levels rule out the use of human labour and as yet undeveloped automatic, remote techniques will have to be used. Indeed, the very possibility of such cracks occurring was not even considered in any of the preliminary safety calculations. The only other possibility is permanent shutdown after five years which, given the present cost of 3000 million FF for each reactor, would render the whole nuclear programme economically unfeasible.

These were the facts made public by both government and unions. Along with their report the unions demanded that the loading of fuel elements at Gravelines, Tricastin and Dampierre be halted until further checks were made. The EDF replied 11 days later, in issuing their report, that the loading would continue on schedule. The CGT and CFDT then called out their members in a five day strike which ended in the government's agreement to a cessation of loading and a new investigation. While this second investigation was underway the Ministry for Industry launched a counter-offensive in the media. On 10 October a ministry official said that the government had instituted its investigation long before the unions' announcement, while government spokesman M Marcel Boiteux said during a television interview that 'where (cracks) have been found they are not serious; where they might be serious they have not

yet been found'.

But further trouble had already arisen. On 10 October there was a leak of primary coolant at one of the Bugey reactors, forcing a shutdown. Now under severe pressure, the EDF announced a blanket censorship on any further information on the nuclear programme; all further statements were to be subject to government clearance. In response a petition which had been circulating, calling for a national debate on nuclear energy and the suspension of all work on nuclear reactors until the conclusion of that debate, rapidly gained sponsors and signatures. The discovery of further leaks at two other Bugey plants and the announcement by the EDF on 24 October of their decision to resume loading at Gravelines (followed by a similar decision for Tricastin though not yet for Dampierre) brought things to a head.

The hitherto united trade union front broke. The CFDT, which like the Socialist Party opposes the present form of the nuclear programme, called for a strike at Gravelines to prevent loading. But the CGT, which along with the Communist Party supports an all-nuclear policy, accepted the EDF decision. Since the CGT was the union of the majority of nuclear technicians the CFDT call was heeded by a comparatively small number of workers. Even so, the loading at Gravelines and Tricastin has had to proceed under police protection with technicians brought in from outside the EDF, an unprecedented situation even in security conscious France.

The impact of the national petition for a national nuclear debate, is not yet known. Public opinion polls before the announcement of the cracks showed pro-nuclear opinion ranging from 53% to 62% with an opposition between 24% and 37%. During recent union elections within EDF, however, the anti-nuclear CFDT gained 2.5% while all other groups, including the pro-nuclear CGT, lost ground. Since this swing now gives the CFDT a majority position at Gravelines and Tricastin, any government decision to start up these reactors after loading is completed in a month's time may be expected to spark off further protest.

Meanwhile, with three of its six operating reactors closed down, the EDF is facing difficult choices this winter. It has already called upon industries to help produce part of their electrical needs themselves, and it cannot avoid unpopular power economies in the domestic sector in the months ahead. With the further delays in the commissioning of new PWRs attendant upon the discovery of structural faults, the vulnerability of an 'all-nuclear' policy has become more evident. □

Jim Ritter is in the department of Mathematics, King's College, London

Development: where are the real experts?

This year's expensive jamboree of United Nations conferences on science and technology for development yielded hollow phrases and little else, writes **Sir Ieuan Maddock**. Funds would be better directed to small teams of trouble-shooters who have direct experience of real problems, such as flood control in Bangladesh or health care in Tanzania; the teams would make "blunt appraisals" and recommend practical actions, building up a library of useful prescriptions.

If noble phrases, robust restatement of the obvious and general amiability of colleagues is the formula for peace on earth and universal prosperity then the symposium organised by the United Nations' Advisory Committee on the Application of Science and Technology to Development (ACAST) must indeed be regarded as a historic event.

"The New World Economic Order", "Immediate and Total Disarmament", "Free and Open Flow of all Scientific Information", "Global Mobilisation of all the World's Scientists" and many more such phrases crackled and thundered around the chandeliers of the Festsaal of Vienna's magnificent Hofburg like a Wagnerian overture. The facts that these noble objectives were very far from being achievable, and that none of the world's scientists had even an inkling that they had been mobilised (by whom? for what?) did not deter the faithful. The existence of the Iron curtain, the troubles of the Middle East, the problems of South East Asia, the stresses in Africa, the conflicts in Central and Southern America, the agonies of Northern Ireland and the economic ills of so many countries were brushed aside. Noble declarations were good for the soul.

Down from Olympus

When ultimately the delegates agreed to clamber down from the top of Mount Olympus they proved to be disappointingly mortal. After turning their collective wisdom to such subjects as food and agriculture, health, natural resources, transport, population, and energy, a series of improvised, working groups produced conclusions which were banal in the extreme. Nothing was said which had not already been said far more competently in a multitude of conferences on the same fashionable subjects many times before. The level of originality can be judged by a few quotes chosen at random as follows:

●On natural resources: "Recommends to government that priorities be set for mineral and energy resources, water resources, marine resources and other natural resources in their development policies". "The multidisciplinary nature

of the concerned science and technology requires technical cross-sectoral mechanisms for cooperation and co-ordination".

●On transport: "That the integrated policy and planning for transport must be evolved in developing countries, taking into consideration population growth and movement, land use, urban and rural development, agriculture and food production, . . . research and development and comparative study of the development of transport on the international scene in order to avoid errors of the past".

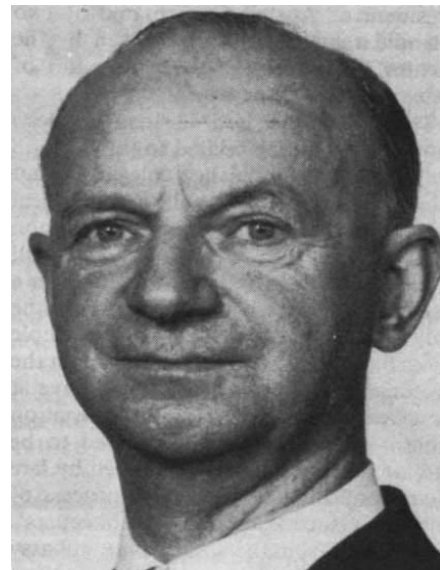
●On communication: "More attention should be given to the value of telecommunications as one of the important instruments for development. The various forms of telecommunications and their usefulness should be investigated. For example radio broadcasts should not be underestimated".

●On population: "Recommend an increase in real terms in the investment in reproductive research and contraceptive development and adoption". "Because population is both a component and beneficiary of development it is recommended that attention be paid to the creation of development models which feature the dynamics of population".

●On human settlements: "Generally there is need for strengthening the scientific, technological and administrative capability for human settlements planning, policies, strategies and implementation to meet the basic needs of the people. This includes re-orientation of architecture, engineering, planning and management of human settlements".

●On environment: "The scientific and technological community is urged to give increased emphasis to interdisciplinarity as an operational concept for tackling the complex land-use problems of ecosystem management".

●On energy: "All possible types of primary energy resources need to be developed including hydrocarbon, coal, nuclear, hydro and solar. When one of these sources is developed even by a few countries it will release non-renewable resources like hydrocarbon for other countries. From this point of view



Sir Ieuan Maddock: the jet-setters must be firmly pushed aside

development of coal and nuclear energy appears to be necessary".

●On industrialisation: "These relate to the enormous technological developments of the developed countries as well as the accumulated problems such as unemployment, low productivity, unequal income distribution in the developing countries, posing major challenges to the technological and industrial developments".

Even the strongest digestion tended to become debilitated on a constant diet of this kind of stuff. If one felt prompted to make a comment based on practical realities, one is inevitably felt a bit of a spoilsport. Why diminish the heady delights of playing God.

I came to the conclusion that the ACAST conference was just about the least effective I had ever encountered — that is until I experienced the United Nations Conference on Science and Technology for Development: UNCSTD.

UNCSTD was held in a giant sports hall with some two hundred individual delegations each with its own table, six delegates' chairs, translation equipment and microphone — the minimum kit, it seems, of any serious minded participating country. Several of the larger nations had 'supporters clubs' seated on the sloping galleries around the floor. Each delegation was allotted fifteen minutes to make a statement and these were interspersed by statements from sundry United Nations organisations and a sprinkling of acknowledged non-governmental organisations (NGOs). After the initial ceremonial opening by the Secretary General of the United Nations and by the

Sir Ieuan Maddock is Secretary of the British Association for the Advancement of Science. He represented the Royal Society at the UNCSTD and ACAST conferences. This article has been submitted to the Council of the Royal Society, and it is now published with the Council's approval.

President of Austria, punctuated by two splendid renderings of Beethoven by the Vienna Philharmonic, the procession of statement-makers began.

The distinguished delegate from Ruritania would be invited to the rostrum (we were all 'distinguished delegates') and each would make an almost identical speech. This process started around noon on the first Monday and was still proceeding late on Tuesday evening over a week later. One wit commented that the only worthwhile statement in the whole series had been made by Beethoven on the opening day. It was obligatory to have at least one person sitting at the delegation table — although the hall tended to be embarrassingly sparsely populated by late afternoon, and so there was a process of finding a succession of 'watchkeepers'. Junior staff from the appropriate embassy or academics who chanced to be studying in Vienna were pressed in service.

Action in the corridors

Inevitably most of the action occurred in the corridors and more particularly in the 'groups'. There was the 'Group of 77' which represented an initial gathering of 77 less developed nations (there were by now over 120). Then there was the EEC, the Eastern European Group, and the Western European and other Countries Group (known as 'WEOC'), an Asian Group, an Islamic Group and so on and so on. These were constantly meeting, determining 'positions', negotiating between groups, and colliding with each other in a manner that could have provided a dozen plots for C P Snow. It was small wonder that the only outcome of this much publicised meeting was to be a weak soup of half-boiled ideas and heavily spiced polemics.

Much of the trouble arose from the skew distribution of the interests and outlook of the people who served in the preparatory phase and also the main conference. At UNCSTD the delegates were drawn almost exclusively from the ranks of academia, of national or international research institutes, of bureaucracies and of politics.

It was a remarkable defect of the structure of the delegation that there were so few people present with any direct experience of the application of technology. Why were names like Ford, Siemens, Ericson, Hoechst, Roche, Heinz, Mitsubishi, Massey-Ferguson, General Electric, Brown Boveri, Unilever, and Shell absent? Where were the hospital superintendents, the forest planters, the dam and bridge builders and sewerage plant constructors? Indeed water and sewerage received hardly a mention throughout the conference, presumably because these were hardly fit subjects for eminent scientists to be found to be discussing. Why were consulting control experts absent from this conference? Where were craft-guilds (the analogue of

London City and Guilds) the technician training boards (such as EITB in the UK) the inspectorate and the specification bodies (BSI in the UK)? Why were there no farmers, fishery managers, food packaging experts, food inspectorates or catering managers to be seen let alone heard? Should not those with direct responsibility for and experience of operating railways, urban roadways, electricity and gas utilities, ports and harbours, fire fighting and ambulance services or telephone networks be at the forefront of the endeavour to harness science and technology for development?

But the whole conference had been conceived and executed by those who so blinded by the pyrotechnics of research that they failed to see, or to admit to seeing that there were others far more relevant and more experienced than themselves for the task in hand.

Although several speakers pointed to the essentially long time scale between even successful research and its ultimate impact on society on a significant scale, the emphasis on research continued. To point out that there was already one billion people on the planet living at the survival threshold and that this number would at least double and possibly treble in the next twenty years — a mere thousand weeks — could not deflect the research enthusiasts from their course. Indeed it was clear that some of the less developed countries thought that, despite the inescapable validity of the points made, the developed countries were raising these points only wilfully to impede the development of the less fortunate countries.

There was the persistent impression that we were dealing with a 'global R&D fund' and that somehow the developed countries had seized 97% of this. The fact that the R&D figure was nothing more than the aggregate of a great mass of individual and unrelated decisions seemed to be difficult to convey. That economic and social order was a prerequisite and not a consequence of a high R&D expenditure was only vaguely understood.

I feel convinced that this intellectual gap would not have been present had the composition of the people attending both ACAST and UNCSTD been different. If the core of the delegates had been drawn from those with direct responsibility for and experience of *applying* technology then the community of purpose would have bridged the gap of semantics or of doctrines. Of the delegates less than 10% could be clearly identified as people with direct industrial responsibility and experience.

The real experts

Alas! the problems of harnessing the fruits of science and technology to meet the needs of the less developed nations remains. It is a sad thought that in the two years that it has taken to prepare the UNCSTD conference at least one hundred

million people have been added to the ranks of the starving and the diseased. It is pointless continuing to hold giant congresses on this or that aspect of the subject and thus providing fodder to that intellectual jet-set who thrive on a menu of agenda drafting and resolution passing. The problem is too urgent and too serious to be left in such hands. All those who control the aid funds (which in total are quite large if they could be efficiently deployed) should cease to support them.

Instead, small teams of *real* experts, experienced in the technical, human and economic aspects of each subject area, should be brought together. If it is flooding in Bangladesh that is to be discussed, let flood control experts supported by economists, and management experts be brought together to make a long visit to Bangladesh to discuss matters with those directly concerned, give seminars, inspect facilities (including education and training) and to understand the local and cultural constraints. Let such teams write precise and, where necessary, blunt appraisals in cooperation with the local experts and authorities and let them prescribe a plan of action. If it is the health problems of Tanzania, the housing of Peru or communications in Lapland, let the same general procedure be followed. In this way a library of prescriptions would be built up from which both specific and general action programmes would be derived. Above all a network of effective human links would be established between those who are responsible at the national or local level and those who have access to the relevant technologies.

\$50 million meeting

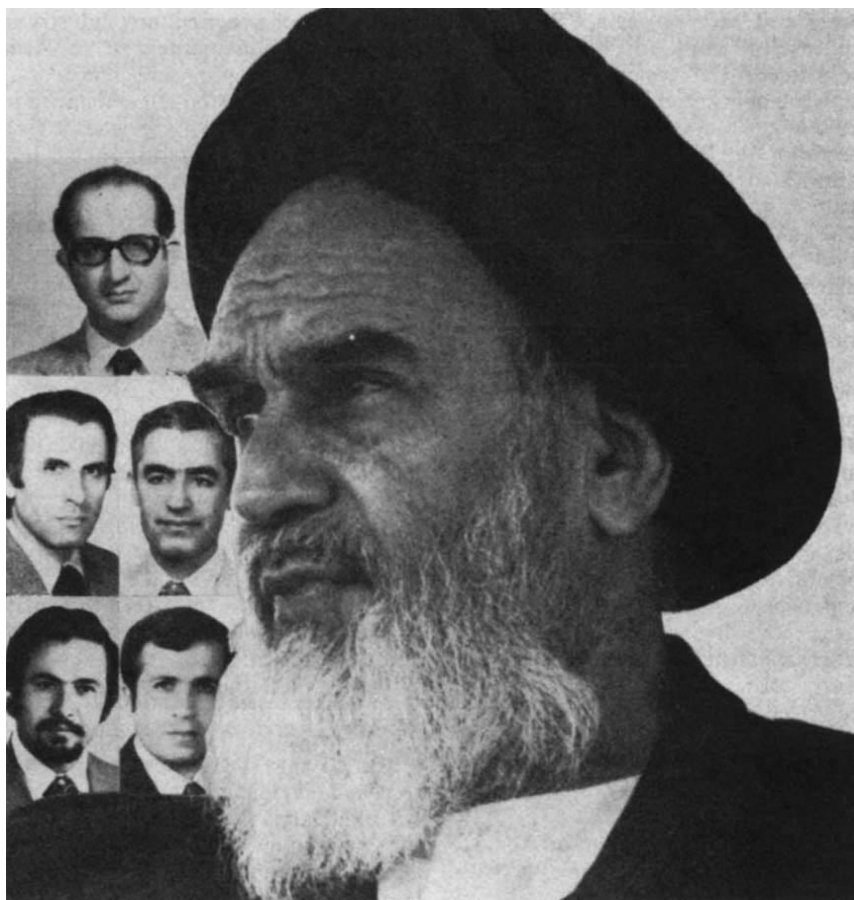
Such a scheme would cost money and take time — but the UNCSTD conference took two years to prepare and is reputed to have cost fifty million dollars. The same amount of money deployed over the same time period in the way described above would already have yielded a great fund of exact knowledge on which current and future programmes would be based. The asset would be accumulating as the mosaic of local pictures built up.

It would be easy just to accept their conference as an agreeable if expensive frolic and to excuse it as the price which has to be paid in the struggle to establish some kind of a world society. Time and circumstances do not permit this kind of tolerant cynicism. With one billion people already at the starvation level and the certainty that this number will treble before the end of the century, it is time for action rather than for conferences. The jet-setters must be firmly pushed aside, they should receive no further encouragement by financial grants or assistance in kind and the task must be steered towards those who are realistic and expeditious. There is no shortage of these if only they are given the opportunity and the support. □

Scientific thinking behind Khomeini

Hidden within what appears to the West to be an increasingly anarchic Iranian revolution lie a number of young thinkers and ideologues who are planning a new science and technology for Iran. Bitterly critical of the past, they foresee a science which would respect basic human needs and Islamic values. **Zia Sardar**, an enthusiast for their cause, spoke to them

Right: Ayatollah Khomeini and Iran's science policy makers from top and l to r: F Darvish, A Safai, M N Bahadori, M Roohi, A Barzegar.



How will Iran use science and technology to achieve its aims? And what is the role of science and technology in the realisation of an Islamic vision? "We take it for granted that science and technology are a means to an end, and not an end in themselves", says Dr Mahdi Nejat-Bahadori, Professor of Mechanical Engineering at the University of Shiraz and a member of the sub-committee of the Revolutionary Council on Planning in Science and Education. "So we must ask: what ends do we seek?"

"What we have agreed to do", says Dr Bahadori, "is to redefine development as a compound of economic, scientific, technological and social actions that take the society closer to God. And this is used as a guiding principle in all our planning". Within this principle, activities that promote development will be encouraged. But certain scientific and technological activities are too complex and their impact on society cannot be easily appreciated. "For this reason", says Dr Bahadori, "we want all societal changes to be slow, after due consideration and with full consent of the society. We will not commit ourselves to a particular science and technology policy without full consideration of its impact on the society, debate and discussion by our people of it and of its place in our development goals."

Dr Bahadori emphasises that the Iranian revolution is not an economic revolution. "We are not aiming to have a prosperous or a consumer society. Or to further the interests of a particular class or group. Our

aim is to elevate human values and our science policy will reflect this aim in all its manifestations".

Dr Ali Barzegar, Deputy Minister of Higher Education argues that a science and technology which elevates human values and the principles of Islam cannot be imported. "They have to be Iranian through and through", he says. "We are not attracted towards big, conspicuous projects with grandiose goals. We are more interested in small, more humble and humane projects that do not get out of control and that utilise our abilities and resources and develop local capability. We want to use science and technology to meet our basic needs and, as these needs are changing, our planning will be continuous. The important thing to realise is that if you do not want to create a consumer society, or chase excessive profits, production becomes a second order determinant. And you can concentrate on quality".

Within the new scheme indigenous research will become very important. Dr M Zirakzadeh, the former deputy minister responsible for scientific research (now replaced by Dr Syed Mohammad Mahdi Jafari), says that universities, government bodies and private industry will be encouraged to tackle local problems and offer viable solutions. Financial incentives will be provided. "But the allocation of funds will be on the basis of projects and their importance to Iran, and not on the basis of organisation and research centres", he says. At present there

are some 120 research centres in Iran but, according to Dr Zirakzadeh, more than half of them have either been transferring technology or using foreign consultants to get their job done. "Most of these", says Dr Zirakzadeh, "will either be closed or reorganised on a more appropriate basis."

The old National Research Council, which before the revolution was a special committee of the Ministry of Science under the chairmanship of Queen Farah, has been retained with a new mandate to reorganise and revitalise local scientific research. Fields singled out include agriculture, cottage industry, public health and alternative sources of energy, and the development of links with industry and research bodies.

While the new Ministry of Science (a reorganisation of the old science ministry) has been working on a new science policy, a sub-committee of the Revolutionary Council has been working on the new Development Plan. This plan marks a radical change from conventional development planning, just as the Iranian definition of development differs sharply from the commonly accepted ideas.

It is a sophisticated plan with a 12 year horizon, split into three phases. After a two year transition period, the plan proper begins to unfold. It will operate within fixed limits but there will be considerable scope for change and adjustment. Each phase will start at a presidential mid-term so that changes in details can only be made when a particular government has been in

power for at least two years. "While there is plenty of scope for rethinking and readjustment, changes cannot be abrupt; and the planning process is continuous and over-riding," says Dr Ataollah Safai, Director of the Centre for National Spatial Planning of the Plan and Budget Organisation.

The plan is based on current predictions of both the needs and requirements of Iran and its resources and potential twelve years after the revolution; and it assumes that during this time oil revenues will decrease considerably. It aims to take Iranian society closer to Islam, and has several strategies running in parallel.

The Development Plan concentrates on four areas. First, the fundamental sector of industry, where the emphasis is put on private small cottage industry. Medium-scale industry will be operated on a cooperative basis with the government and essential heavy industry will be run entirely by the government. "The economy is neither capitalistic nor socialistic", says Dr Safai. "Private industry will be encouraged but monopoly will not be allowed. The state will not interfere with the freedom of the individual nor will it have total control over the means of production."

Second, considerable attention is paid to the distribution of economic activity. Strategies for de-urbanisation and rural development are recommended, and a balance in economic activity is sought between the national, regional and local levels. There will be separate programmes for urban, rural and regional systems.

The third area covered is the development of traditional economic

sectors — such as agriculture, industry and services within the context of an Islamic society. Both agriculture and cottage industry are given equal emphasis, but services are limited to a level where they do not interfere with the spiritual and cultural values of society. According to Dr Bahadori, the plan will meet all the basic needs of the citizen while at the same time checking materialism. Finally there is an area covering the environment and the creation of a capacity for future expansion, which are intended to create a physical and spiritual environment conducive to the Islamic way of life.

In the area of science policy, the plan emphasises creativity and originality. Dr Safai says "an environment that encourages thinking, with maximum freedom and inducement for originality and creativity — this is our survival imperative". The plan condemns blind imitation of the Western scientific establishment, and envisages the expansion and reorganisation of the educational system to meet the educational needs of Iran.

For education, the country will be organised into regions which will then receive individual attention. This, says Dr Farrokh Darvish, Director of the Institute for Planning in Science and Education, is already happening. Education in Iran has been mainly a paper chase for better jobs and promotions. "Now we are emphasising education which is geared to regional needs", he says.

Although the Iranian educational system faces many new problems such as an acute shortage of teachers, administrators and text books (most curricula have been

changed in content or in emphasis) most of its problems are typical of the Third World: irrelevance to social needs, over-emphasis on higher education and on paper qualification.

Although the new Development Plan has been finalised, it awaits presentation to and approval by the Revolutionary Council. Pending its start, the government has cancelled all technology transfer projects, the nuclear energy programme and heavy industry schemes and has reduced the budgets of the existing educational, scientific and research organisations and frozen all expansions that were planned before the revolution. This has created a certain amount of temporary unemployment and ill feeling in industry and education. But, according to Dr Barzagar there is little else that can be done until the society has direction and the plan comes into operation.

The policies put forward so far in the pre-plan period, have been mostly negative however, and some of them have created an uneasy atmosphere. There is also a real danger of over-reaction to the Shah's extreme policies. The pool of technical manpower created under the Shah is now a major asset which must not be forced into self-exile by over-zealous revenge seekers. The revolutionary leadership should try to preserve all of Iran's human resources.

Most of all, the revolutionary leadership must ensure that balance and clear thinking dominates its policies: zeal can sometimes blind reason and humanity. If Iran can avoid this trap it could become a model for all those who seek to build a society on human and spiritual grounds. □

The multinationals come under scrutiny

THERE is considerable ill feeling in Iran against the American and European multinationals and business concerns. "They have plundered Iran without any ethical or moral consideration", says Dr Barzegar. The sales techniques used in Iran, and no doubt they are now being used in other Third World countries, were a balanced compound of Skinnerian Psychology and gangsterism, he says. "Most of the contracts signed with the multinationals by the Shah are ruthlessly one sided".

There are many examples of such joint ventures and technology transfer agreements. In one case, the Shah agreed that General Motors should set up a subsidiary in Iran with the parent company holding 45% of the capital in the US. The agreement stipulated that all key managerial positions, including managing director, treasurer, sales manager and supply manager would be held by American citizens for an indefinite period and that the US parent company would charge the Iranian company a managerial fee and an engineering fee equal to 15% of the cost of machinery.

Even the track record of Scandinavian

companies (Scandinavia is considered to have liberal policies towards Third World problems) is poor. For example, consider the case of the Iranian Ball-bearing Company, established in 1969 as an \$8.6 million joint venture with the Swedish firm SKF. Even when the Iranian plant began production, SKF continued to import ball-bearings into Iran. As the local product lacked quality, the Iranian Ball-bearing Company suffered tremendous losses and had to reduce its capital by half. SKF, however, made substantial profits.

Iran's dealings with Electronic Data Systems (EDS) are also notorious. Under the Shah, an imaginary demand for sophisticated computers was created. For example, Dr Shojaeddin Sheikholeslamzadeh, Minister of Health and Welfare under the Shah, was trying to implement a mad drive "to have computer terminals in every village in the country, even the remotest, which will let us know immediately a person is ill, so that a central computer can notify the authorities who will send a helicopter to transfer the patient to the nearest hospital". Similar logic was being followed in the Plan and Budget

Organisation, the Social Security Organisation, Iran Air, and other government as well as private organisations.

The EDS signed a contract with the Social Security Organisation to computerise its operations. The company supplied a number of packages, without the necessary manpower. None of the packages, however, could meet the needs of the Social Security Organisation. EDS then asked for its contract to be extended from the original three years to five years and the extension was agreed. After four years and \$80 million paid out, it became clear to the Shah's regime that EDS had come to stay permanently. Shortly after the firm quadrupled its original demand, with the project still less than one-third complete, Iran had its revolution.

The firm has now sued the Iranian government for breach of contract and the remainder of its money. The action has been taken in the United States and EDS has succeeded in securing the confiscation of the Iranian government's property in the US.

The nuclear programme of Iran under

the Shah is another example where the Iranians feel they have been exploited by the multinationals. The programme envisaged a network of 23 reactors, supplying 42% of Iran's electricity needs by 1994. Before the revolution, work on four nuclear plants had started and a further 19 were being negotiated.

According to Mr Mansoor Roohi, Vice-President of the Iranian Atomic Energy Organisation (IAEO) the contracts signed for the four nuclear power plants, two to be built by the German company Kraftwerkunion, an engineering unit of Siemens, and two by the French company Framatome, were "remarkably one sided". He claims that the IAEO were warned not to interfere with the work of the foreign companies in case they used this as an excuse for delaying the work. Framatome charged \$2.3 billion to build two pressurised water reactors of 900 megawatt near Ahwaz on the Karun river and Kraftwerkunion charged \$3.2 billion for two plants of 1,200 megawatts near Bushehr on the Persian Gulf.

According to the IAEO, in order to meet the demands of its nuclear programme, the Iranian government purchased vast quantities of uranium at inflated prices,

and directed it to Germany for enrichment. Once enriched, the uranium had to be stored in Germany till 1994, or until it would be needed by Iran. This had been going on since 1974. The Iranian government paid Germany both for enriching its uranium and storing it.

By 1978 the cost of both projects had more than doubled. 60-70% of work on the German plants had been completed, while the sites for the French plants were still being prepared. When regular payments stopped during the revolution, both companies pulled out. They are now threatening legal action.

Iran has decided not to continue with construction of the four nuclear plants and with plans for the remaining 19. A 50-man special commission concluded that the German company had either failed to fulfill all its undertakings or had done shoddy work.

For example, 400 Iranians were sent to be trained in Germany. The commission claimed that they were given the most rudimentary training not much above the technician level. It also considered that the \$7 billion sum demanded by the German company to complete the project was

unrealistic and that even the \$3.3 billion already paid out was grossly unfair for the work done. Sadeq Tabataba'i, Deputy Prime Minister in the Bazargan government, made it clear that "Iran has no intention of paying the compensation".

Meanwhile, the IAEO is rethinking the long-term energy policy of Iran. "We had to abandon the nuclear energy programme because it was designed to keep us in the hands of the foreign companies for decades. Although we are looking into other sources of energy such as solar energy and hydro-electric power, we have not completely ruled out the nuclear option", says Roohi.

Iran is continuing some of its nuclear research and training programmes and a small power plant of 20-30 megawatts will be kept in operation, so that it can retain some nuclear option for the next 10-12 years by maintaining a certain amount of indigenous capability and manpower. However, Roohi thinks that for the "next decade or so, the time needed for us to explore other alternatives and come up with a new, more appropriate energy strategy, Iran will rely largely on its oil to meet its energy needs." □

Iran faces an environmental crisis

Dr Taghi Ebtekar, the newly appointed director of the Department of Environment, says that urgent action is necessary to stop further despoilation of the environment. According to Dr Ebtekar, the Shah's economic policies of rapid industrialisation without any consideration for the environment, as well as rapid urbanisation and a growing population, have resulted in alterations to and destruction of many ecosystems.

Complete river systems, such as those of the rivers Zandron and Anzali (formerly Phalavi) in the Caspian, are under severe threat from industrial wastes and emissions. Hard detergents from industrial plants have done serious damage to marine life. The Caspian sea is now heavily polluted with "dangerously high levels of DDT, mercury and other heavy metals". "A recent study carried out after the revolution in the northern part of Iran, reveals that the level of DDT and mercury in mothers' milk is up to five times higher than the accepted critical level", says Ebtekar. The Caspian forests, "unique among the temperate forests of the world" are also in great danger.

Vegetation in the northern part of Tehran, he says, is under heavily polluted water and in Ferozabad natural vegetation is swamped by sewage. Air in Tehran contains up to four times the permitted levels of particulates. In Shiraz and Isfahan, air pollution is attacking cultural monuments and threatening the world renowned Iranian architectural heritage.

Soil erosion and soil pollution are also major worries for the department. In the steppes of western Asia, for example,



Traffic pollution is a major problem.

enormous areas are turning to desert. The situation has become so serious, according to Ebtekar, that unless measures are taken to halt and reverse the process, there will be total loss of vegetation over wide sectors with consequent erosion, floods and dust-storms. In other areas, soil, a scarce and precious commodity in Iran, is being polluted by pesticides, heavy metals and fertilisers.

Ebtekar blames the Shah's economic policies for much of the urban pollution. For example, "the Shah promoted private transport and ignored public transport. As a result we have almost 1.2 million cars in a city of 5 million people. Over 80% of urban pollution is the result of exhaust fumes". The Department of Environment has persuaded the government to introduce new, more comfortable buses and to expand the bus routes. It is planning legislation to limit the use of private vehicles and heavily leaded and poor quality petrol. Iranian vehicles will also require "certain devices" to eliminate hydro-carbon and carbon monoxide fumes.

Legislation is also being prepared to enforce appropriate waste-treatment

facilities in polluting industries. Sets of National Emission Standards have been developed for application to most industries throughout the country. The department also plans a wholesale conversion to natural gas both for industrial and domestic fuel. Conversion had already started in Tehran before the revolution and will soon be extended to Shiraz and Isfahan.

The department administers 62 reserves for protection of the floral and faunal resources of Iran, constituting one of the most comprehensive programmes for nature conservation in Asia and the Middle East. A number of endangered species are protected in accordance with the wildlife regulations. Of the animals in this category, only the Caspian Tiger (*Felis tigris*) is genuinely endangered, possibly being extinct in Iran. A number of other species, such as Asiatic Black Bear (*Selenarctos thibetanus*) and two species of gazelle and Onager, once believed to be extinct, are now considered safe.

Conservation and protection of wildlife and natural environment are dictated by Islamic law, says Ebtekar. As such, conservation is being promoted at the popular level. "We are conserving our natural resources for the people", he says. "The people must have access to our wildlife, wetlands and other natural resources. We do not want to isolate conservation areas and preserve them as museum objects. Neither will we limit access to conservation areas to certain privileged classes. Everyone will have access to conservation areas as long as conservation laws are obeyed". □

correspondence

Three Mile Island not "a major disaster"

SIR, — I am writing to comment on the article in your issue of (8 November, page 120) on the Three Mile Island report. In the heading you refer to the accident as "a major disaster". In the text you write "the worst accident in the history of the US nuclear programme". The reader might think there must have been a large number of people killed and injured. In fact there were no deaths and no injuries except psychological. Estimates of the long-term consequences have suggested that a few (less than 10) additional cancers may ultimately develop in the population exposed, an effect likely to be far outweighed by the cancers caused by an extra packet of cigarettes smoked by the same population, or the additional traffic accidents due to the upheaval. How about deaths from bronchitis due to pollution of the air from coal-fired power stations and factories?

A journal of the standing of *Nature* should not be a vehicle of sensational reporting. Many of the recent comments in your columns have shown a total lack of sense of proportion in that the dangers of radiation are never put in the context of the risks we all accept in daily life, nor of the risks attached to not using nuclear power. An irrational fear of radiation has been built up by the popular press and other media; serious journals such as *Nature* should take a more rational and responsible attitude.

Yours faithfully,

D.K. BEWLEY

Medical Research Council, London, UK.

Scurrilous comments belie Egyptian-Israeli cooperation

SIR, — It was difficult to understand the relevance of Dr Galal's comments (11 October, page 424) on the article by Z. Sardar (2 August, page 350). The hostile views expressed by the author were in complete defiance of the brave and new approach of Egypt's President Sadat. In spite of Dr Galal's comments I hope that not all the scientific community in Egypt share his views.

It seems that during the twenty years he has been science editor of *Al Ahrām*, Dr Galal has adopted the political approaches, but has failed to overcome the psychological gap to which President Sadat has referred many times. What is the relevance of Dr Galal's remarks on "Israeli arrogance in science and technology" to the potential benefit to scientific cooperation? Where did he get his impression of the "superiority complex" of Israeli scientists; or where did he get the idea that Israeli scientists regard their Egyptian colleagues as incompetents who need teaching by us?

I would like to remind Dr Galal that Israeli science has suffered no less than Egyptian science from the continuous wars, and his scurrilous statement on the "colossal funds and privileges that Israel is getting from rich Jewish communities" is far from being "objective, realistic and frank". The money that was given to Israel has been spent to absorb hundreds of thousands of refugees from all over the world, mainly from Arab

countries. Was the money from rich oil-producing Arab countries used to improve the condition of the Palestinian refugees? They were kept for more than twenty years in the most miserable conditions and used as a political football. Is that not depriving the Palestinians of their right to live as human beings? The Palestinian problem will be solved only when the Palestinian policy makers stop aiming at the destruction of Israel and recognise the right of the Israelis to live in their homeland. I hope that the Egyptian people, including the scientific community, now agree with their President and recognise the State of Israel, in spite of the problems that still remain to be solved.

Were this to be so, I feel sure that true cooperation between scientists from Israel, Egypt and the world can be achieved, to the benefit of all the peoples of the area.

Yours faithfully,

Z. BEN-ISHAI

Technion, Haifa, Israel.

Antijewishism

SIR, — The style of the letter from Mr Salah Galal (11 October, page 424) is perhaps appropriate for *Al Ahrām*, but for *Nature* it seems to be unusual. Though some of the Israeli scientists may have superiority complexes or may be arrogant, the supposition of a national superiority complex and arrogance of Israeli scientists in general is nothing else than antijewishism. Thus, disregarding their demagogic content, the arguments and statements of Mr Galal have a diminished value because of their motivation by his obviously inherent antijew emotions.

Yours faithfully,

LAJOS ERNST

Freie Universität Berlin, West Germany.

Distorted interviews in East Germany

SIR, — When I attended the meeting of the Federation of European Biochemical Societies (FEBS) in July 1978 in Dresden, I was invited to give an interview to the Press. This was conducted in German and recorded.

Now, in November 1979 I am surprised to receive "with thanks for my collaboration" a book with contributions to mark the thirtieth anniversary of the GDR (*Of the Birth and Growth of the GDR, Zeit im Bild, Dresden*) in which my question-and-answer interview has been included in the form of a 500-word statement in English. The translation has resulted in a eulogy of the GDR and of GDR scientists, which in its exaggeration is ludicrous. The poetic licence of the translation has caused subtle, and some not so subtle, changes; but when — to give but one example — I am quoted as referring to "a secure existence free from worries of unemployment", I am absolutely certain that I said that I was not competent to answer questions on full employment in the GDR.

I do not blame the colleagues who organised the FEBS meeting in Dresden, of which I still retain a happy memory, but would counsel visitors to the GDR to learn from my experience.

Yours faithfully,

H LEHMANN

University of Cambridge, UK.

Health care in India

SIR, — Anil Agarwal's article on 'barefoot doctors' in India (30 August, page 716) creates the impression that a new health policy was launched in April 1977 after the Janata Government took over. In fact, the community health worker (CHW) made his appearance in the planning of the Ministry of Health of the government of India much earlier. In 1972, he was called the Rural Medical Practitioner (RMP). The previous government had hoped to deliver a modicum of health care to the entire country, using RMPs, by 1976. It is apparent that the goal was never reached.

The problem in India at present is not that of an inadequate supply of doctors; it is that of maldistribution. The reason for this geographical lopsidedness in the distribution of health care personnel is that recommendations of the Bhore and Mudlair committees were never fully implemented. Professional isolation, lack of adequate facilities in the practice of scientific medicine and often a lack of personal security are a result of this apathy on the part of successive governments in India in implementing the recommendations of its own specialised committees.

To answer the question with which the article ends: yes, there are alternatives to the CHW scheme. Building physical facilities for health care in the rural areas (as was recommended more than two decades ago) may take another two decades to accomplish. In the meantime, doctors could function efficiently if organised along the lines of the Indian Administrative Service or the Indian Police Service.

The junior-most officers of these cadres very often live in towns, where all the facilities they are used to are easily available, and commute to their places of work in the different villages. Doctors in an Indian Medical Service could do the same.

The children of these doctors could then attend schools which their parents think are reasonable, thereby eliminating one major reason for their reluctance to work in rural areas. Professional isolation would be done away with at the same time, as they could consult with each other on returning to the town after the day's work. Coverage at night could be provided on a rotational basis, akin to the duty hours spent as house-officers in urban hospitals. All this of course implies government expenditure.

Another alternative is completely outside the government. Such a scheme would draw on the private sector — the mini-international companies which Anil Agarwal has described in another of his articles (23 August, page 625).

It has been shown to work on an experimental basis, as the Social Work and Research Centre, at the village of Tilonia in Rajasthan. Why not use it as a model for the other parts of the country as well? Contrary to Agarwal's conclusion, one would hope that CHW scheme or the RMP scheme, call it what you will, does die a speedy death. If not, such half baked medical care will bring untold misery to a vast and gullible public in rural India.

Yours faithfully,

BEHRAM PASTAKIA

Michael Reese Hospital and Medical Centre, Chicago, US.

news and views

Production and respiration in animal communities

from Robert M. May

A WIDELY-quoted generalisation about natural ecosystems is that the efficiency of energy transfer from one trophic level to the next is around 10%; that is, about 10% of the net production of plants ends up as net production of herbivores, about 10% of this makes its way into net production of the first level of carnivores, and so on. The generalisation is largely based on studies of freshwater lakes and of laboratory aquaria, conducted in the 1950s. One early articulation of this notion is in Slobodkin's elegant and influential *Growth and Regulation of Animal Populations* (Holt, Rinehart and Winston; 1961), where his speculative list of candidates for valid ecological generalities leads off with "Food-chain efficiencies and ecological efficiencies in nature are approximately constant for all species".

Unfortunately, subsequent research on terrestrial and on other kinds of aquatic communities has overthrown this appealing generalisation, showing that the efficiency of energy transfer from one trophic level to the next can vary very widely. This is a pity, for valid 'ecological laws' are thin on the ground, and we can ill afford to lose any of the few we thought we had.

To determine overall efficiencies of energy transfer, two questions must be answered. First, what fraction of the net production at one trophic level is actually assimilated by creatures at the next level? Second, how do these creatures apportion the assimilated energy between net production (growth and reproduction) and respiration (maintenance costs)? The second question is amenable to fairly precise answers, but the first question is messier, as it can involve both particularities about the fraction of material that is assimilated rather than excreted by a given species and generalities about the overall fraction of net production at one level that is actually used (consumed) by the next level. Some of these problems and ambiguities can be made more explicit by considering, say, mice and weasels. If we focus on the weasels, it is in

principle straightforward to determine the efficiency with which 1 gram or 1 calorie of mouse eaten is transformed into grams or calories of weasel. If we focus on the mouse population, it is hard to determine what fraction of their total biomass appears as net production in the next trophic level. Indeed, the answer ultimately depends on how we keep the books; the very notion of 'trophic level' does not stand up to close examination (where, for example, do the decomposers belong?).

Humphreys (*J. Anim. Ecol.* **48**, 427; 1979) has recently drawn together a large body of literature, to determine the relationship between annual production P and respiration R in natural populations of animals. This synoptic study sheds light on the general issues discussed above. It also has more immediate practical applications. If a relation between P and R can be confidently established for a given group of organisms, then for a species in this group only one of the two quantities need be measured directly, which can be a helpful short-cut in compiling energy budgets for communities; several people have used the earlier study by McNeill and Lawton (*Nature* **225**, 472; 1970) for this purpose. Humphreys analyses a total of 235 energy

budgets culled from the literature, and he emphasises that a great variety of different assumptions and possible biases have gone into the individual studies.

Examining regression relations between P and R , Humphreys shows that homeotherms (loosely, warm-blooded animals) can be separated into four significantly different groups: insectivores; birds; small mammal communities; and other mammals. Poikilotherms (cold-blooded animals) separate into three groups: fish and social insects; non-insect invertebrates; and non-social insects. The invertebrate groups further permit significant separation into trophic categories of herbivores, carnivores and detritus feeders. In no case is the relation between P and R significantly different from a simple linear one. The fruits of Humphreys' analysis are summarised in Table 1, which shows the mean 'production efficiency' $P/(P+R)$ or fraction of assimilated energy ($A = P+R$) that is devoted to net production, for the various groupings.

Several interesting patterns emerge. Both for non-insect invertebrates and insects other than social insects, the production efficiency is significantly lower for herbivores than for carnivores and

Table 1 Mean production efficiency, $P/(P+R)$, for various groups of animals.

Group	Mean production efficiency (per cent)	Sample size
Insectivores	0.9	6
Birds	1.3	9
Small mammal communities	1.5	8
Other mammals	3.1	56
Fish and social insects	10	22
Non-insect invertebrates	25	73
Non-social insects	41	61
Non-insect invertebrates		
herbivores	21	15
carnivores	28	11
detritivores	36	23
Non-social insects		
herbivores	39	49
detritivores	47	6
carnivores	56	5

Robert M. May is Class of 1877 Professor of Zoology at Princeton University.

Table 2. The assimilation efficiencies, or percentage of plant production consumed by feeding-animal species, for various systems.

Plant	Consumers	Percentage of productivity consumed
Beech trees	Invertebrates	8.0
Oak trees	Invertebrates	10.6
Maple-beech trees	Invertebrates	6.6
Maple-beech trees	Invertebrates	5.9
Tulip-poplar trees	Invertebrates	5.6
Grass + forbs	Invertebrates	4-20
Grass + forbs	Invertebrates	<0.5
Alfalfa	Invertebrates	2.5
<i>Sericea lespedeza</i>	Invertebrates	1.0
Grass	Invertebrates	9.6
Aquatic plants	Bivalves	11.0
Aquatic plants	Herbivorous animals	18.9
Algae	Zooplankton	25.0
Phytoplankton	Zooplankton	40.0
Marsh grass	Invertebrates	7.0
Marsh grass	Invertebrates	4.6
Meadow plants	Invertebrates	14.0
Sedge grass	Invertebrates	8.0

detritus feeders. A plausible explanation is that biochemical conversion efficiencies are higher for animals eating other animals than for animals eating plants. Other patterns are shown by Humphreys to be conspicuous by their absence: there is no significant correlation between production efficiency and the magnitude of production (that is, no correlation between P/R and P or R); there is no correlation between production efficiency and animal weight; and, with the groups set out in Table 1, species with different habitats (aquatic and terrestrial) do not have significantly different production efficiencies. Humphreys makes the further point that "there is no quantum jump in production efficiency between poikilothermic and homeothermic animals", but I think Table 1 suggests such a distinction is real (with homeothermic production efficiency typically in the range 1-3%, poikilothermic in the range 10-40%). Admittedly the scatter around the mean values for a given group is such that some social insect species have production efficiencies lower than some mammal species, so that there is no 'quantum jump' between homeotherms and poikilotherms, but the tendency for the typical poikilotherm to have a production efficiency an order of magnitude larger than that of the typical homeotherm remains. Warm-blooded beasts pay a noticeable cost, relative to cold-blooded ones, in order to keep their metabolic machinery ticking over at a constant temperature.

For a community of interacting species,

we can get some idea of the overall 'food-chain efficiency' with which energy flows from one trophic level (n) to the next ($n+1$), by combining the mean production efficiencies at level $n+1$ (P_{n+1}/A_{n+1}) with estimates of the fraction of the productivity at level n that actually is consumed (the assimilation efficiency, A_{n+1}/P_n). Pimentel, Levin and Soans (*Ecology* 56, 381; 1975) have brought together several rough estimates of the percentage of plant production that is consumed by the animal species that feed upon it; their compilation is summarised in Table 2. As mentioned above, any such estimates of assimilation efficiencies suffer, *inter alia*, from the arbitrariness inherent in a crude classification into 'trophic levels'. Convolving Table 1 with Table 2, we see that food-chain efficiencies can vary over two or more orders of magnitude, from less than 0.1% to more than 10%.

In the early 1960s, the tentative '10% rule' engendered enthusiasm for ecological generalisations. The subsequent collapse of the rule, giving way to the complicated variety of patterns shown in Table 1, has, in my opinion, led to an excessive disenchantment with such generalisations. I think the time is ripe to return to these questions, trying to understand the patterns documented by Humphreys and others, both from 'below' (in terms of thermodynamic constraints on production efficiency in different kinds of animals) and from 'above' (in terms, for example, of the possible constraints that energy flow may impose on food web structure). □

Order in amorphous polymers

from Paul Calvert

THE extent of order in the amorphous state of polymers has been debated for many years. Likening the polymer molecules to strands of spaghetti it is difficult to see how they can be packed to high densities unless the strands are largely arranged parallel to their neighbours. After the Chemical Society's recent Faraday Discussion* on Organisation of Macromolecules in the Condensed Phase the situation seems much clearer than two years ago (see *News and Views*, 271, 507; 1978).

Two types of order have been postulated for amorphous and glassy polymers, the orientational correlations just mentioned and local density fluctuations with regions of tight packing separated by a less dense matrix or by boundary zones. The latter were apparently demonstrated by Yeh and Geil about 10 years ago. They observed 2.5 nm nodular structures in many glassy polymers with electron microscopy in bright and dark field as well as in fracture surface replicas. At the recent meeting D.R. Uhlmann (Massachusetts Institute of Technology) reviewed his small angle X-ray scattering (SAXS) results which show only enough scattering to be consistent with the small thermal fluctuations in density, frozen in at the glass transition temperature in polymethylmethacrylate, polyethylene terephthalate, polycarbonate, polyvinyl chloride and polystyrene (PMMA, PET, PC, PVC, PS). This does not eliminate the possibility of heterogeneities whose density differs significantly from the bulk density but they must be present in very small quantities. Uhlmann also concluded that the bright and dark field electron microscope observations of small nodules were due to electron diffraction effects.

Thus the 'typical' amorphous polymers are essentially homogeneous. Epoxy resins are not but SAXS cannot distinguish between small quantities of voids and larger quantities of low or high density nodules. G.C. Stevens (CEGB, Leatherhead) said that small aggregates could be detected in unreacted liquid epoxy resins by light scattering. P.H. Geil (Case Western Reserve University) pointed out that he had seen annealing of nodular structures in amorphous polyethylene (PE) and in plasticised and unplasticised PVC. It does seem reasonable that crystallisable polymers such as PE should form crystalline nodules at low temperatures. PVC is frequently partly crystalline and its structure is most dependent on the polymerisation conditions, so any strange behaviour is plausible, and important, in this polymer.

Thus there are heterogeneities in some systems and possibly in small amounts in

*Held at the University of Cambridge, 25-27 September, 1979.

'typical' polymer glasses, which could still affect the fracture and optical properties. The nodules often seen on fracture surfaces seem to be an artefact.

Most of the evidence presented at the discussion was against orientational order but most of it was on PE and paraffins which are atypical in that the chains are very flexible. E.W. Fischer (University of Mainz) and P.J. Flory (Stanford University) reviewed a range of techniques which point to a lack of regular order. Small angle neutron scattering shows that chains have the same radius of gyration in the glass and melt as for the ideal random walk found in θ -solvent. This result seems quite general, having now been seen for PS, PMMA, PE, polypropylene and polyoxymethylene. Orientational order would be expected to lead to chain expansion, whereas if anything this random coil is a rather contracted state. Infrared and Raman results show that probabilities of *trans* and *gauche* conformations are similar in the glassy, liquid and solution states.

Magnetic birefringence and depolarised light scattering, new techniques in this context, both suggest that chains are not parallel for more than about three carbon atoms as opposed to the strong orientation

Paul Calvert is a lecturer in the School of Molecular Sciences, University of Sussex.

seen in liquid crystals. According to Fischer polystyrene is similar to polyethylene in this respect. A.H. Windle (University of Cambridge) described the impossibility of fitting wide angle X-ray scattering data using models with straight, parallel polyethylene chains where random arrays of straight chains of four or five carbons worked quite well. On the other hand for polytetrafluoroethylene, a very stiff chain polymer, bundle models worked better. These results used scattering beyond $3(\text{\AA})^{-1}$ which are mostly attributable to interactions of neighbouring groups within chains and so are not really a reliable guide to packing between chains.

The weakness of the available data was demonstrated, perhaps unintentionally, by W. Pechold (University of Ulm). His meander model, in which bundles of chains fold into cubic arrays, came in for much criticism and is unrealistic, but does fit most of the available data. Also no alternatives were presented in such detail; most workers prefer to avoid committing themselves on the exact nature of the packing. Thus the chain is random in terms of its radius of gyration and the local bends and curves are the same in glasses as in solution. As for how they actually fit together we can either look more seriously at spaghetti or ignore the question as unimportant. □

Institute of Genetics, Mishima, Japan) reported that a 245-base-pair sequence covering *ori-C*, was the minimum sequence sufficient to rescue ColE1 in similar conditions. Hirota also announced that alteration of even a single base by directed mutagenesis was sufficient to abolish biological activity of the sequence and further studies of this kind should greatly illuminate structure and functional relationships in the *ori*-region.

The presumed significance of the secondary structure identified in all *ori*-regions so far cloned was constantly reiterated and D. Smith (University of California, San Diego) presented new data showing the impressive conservation of possible secondary structures but not necessarily of precise base sequence when the origin regions of *E. coli* and *S. typhimurium* are compared. Overall comparison revealed 88% of conserved bases, and one contributory factor suggested to explain conservation was the presence of 17 GATC sequences, the site of methylation by the *dam* gene, which could therefore be particularly prone to mismatch repair of any incorrect bases in newly synthesised daughter strands.

Some controversy remains over the precise role of the *dnaA* product in the temporal sequence of initiation in *E. coli*. T. Atlung (Institute of Microbiology, Copenhagen) reviewed evidence that the *dnaA* gene product (now convincingly demonstrated to have a subunit molecular weight of 54,000) is an autoregulated, rho-like protein which controls the frequency of initiation, at the primary level, by modulating the rate of transcription in the *ori-C* region. E. Orr (University of Leicester) agreed that the *dnaA* gene product affects transcription but presented evidence in favour of its action at a later step in initiation, subsequent to the primary control step operating in wild-type cells.

Concerning the role of a specific RNA primer at an early stage in initiation, Messer reported that the so-called 'ori-RNA' described previously in his laboratory, does not hybridise with DNA from or near the origin region. Despite this setback there was still general agreement that transcription of a specific ori-RNA primer and/or transcriptional activation to open up the *ori* region, must be an essential early step in initiation. Perhaps indicating the need to unwind the *ori* region, several reports suggested an important role for the enzyme DNA gyrase in initiation. Orr *et al.* and N. Fairweather *et al.* (University of Leicester) described a thermosensitive gyrase mutation mapping in the *cou* (*gyrB*) gene which prevents colony formation. Nuclear organisation, initiation of DNA replication and placement of septa are all apparently defective in the mutant although, intriguingly, the rate of DNA chain elongation is not significantly affected,

Barry Holland is Reader in Genetics, University of Leicester.

Replication dissected

from Barry Holland

At a recent meeting on the control of replication of bacterial chromosomes and plasmids* a mass of genetic evidence was presented that initiation of DNA replication of several plasmids is negatively controlled — that is, through removal of repressor. Even in the case of the high copy number plasmid R6K, where it is known that a plasmid-determined protein (π) is required for initiation both *in vivo* and *in vitro* it was argued that synthesis of this protein is also under autoregulatory negative control (R. Kolter, University of San Diego).

Although the genetic evidence for the repressor is now overwhelming its nature remains elusive. Several participants reported that a DNA fragment coding for a *trans*-acting control substance has been cloned from a number of plasmids but *in vitro* protein synthesis has not yet identified the regulatory substance. This may simply reflect the very small amounts of repressor synthesised, due in part to tight autoregulatory control.

In the past there has been some question of whether plasmid size is a primary determinant of copy number, as plasmids with high copy numbers tend to be smaller. However, there was agreement at the

meeting that, providing the primary control of initiation is intact, increasing the molecular size of the plasmid does not in itself affect copy number; however, if plasmid copy number rises substantially in so-called 'copy mutants' copy number is inversely proportional to size and this indicates that total plasmid DNA then becomes the limiting factor (B. Polisky, Indiana University). Discussion of the multiplicity of potential and active origins found in naturally occurring replicons (for example in *B. subtilis*, F and R6K plasmids) exposed a difference in emphasis between participants concerned with understanding what Nördstrom called the 'basic replicon' (the smallest fragment of self-replicating DNA which maintains the copy control of the wild-type replicon), and others, concerned more with the identification of specific sequences which identify regions of chromosomes capable of interacting with replication complexes to promote initiation. As a step towards defining such regions on the *E. coli* chromosome, W. Messer (Max-Planck-Institut, Berlin) has found a 100 base pair fragment of the origin region (*ori-C*) which when integrated into plasmid ColE1 will drive its replication in mutants in which ColE1 replication does not normally occur. On the other hand, Y. Hirota (National

*The 2nd EMBO Conference on the control of replication and the partition of bacterial chromosomes and plasmids was held on 20-24 September in Leicester and organised by Kurt Nördstrom and Bob Pritchard.

thereby raising doubts about previous assumptions that this enzyme is essential for replication fork movement.

The long haul towards the characterisation of specific components involved in the association of *ori-C* with the cell envelope took an important step forward with the identification of a 65,000 molecular weight membrane protein by A. Jacq *et al.* (University of Paris) and H. Lother *et al.* (Max-Planck-Institut, Berlin) which binds to at least two specific single-stranded DNA sites on the cloned *ori-C* region. Although the precise localisation of this protein in the envelope remains to be established, other findings (H. Wolf-Watz, Umea University; M. Schaechter, Tufts University) that different outer membrane proteins are apparently present in isolated origin-membrane complexes continues to complicate ideas on the actual organisation of such complexes. N. Sueoka (University of Colorado) reported that mutations in a *B. subtilis* gene *dnaB1* (coding for a 35,000 molecular weight membrane protein) block initiation at high temperature, an event which seems to coincide with the loss of chromosomal origins (or plasmids formed from cloned origins) from the membrane fraction both *in vivo* and *in vitro*.

The problem of the relationship between the phenomena of incompatibility of related plasmids and the replication functions of such plasmids was clarified considerably by the general agreement that incompatibility is due to the existence of a *trans*-acting mechanism controlling copy number which fixes the average number of plasmid copies per cell. However, although incompatibility based on competition for a few membrane sites as suggested originally by Jacob, Brenner and Cuzin, was not generally accepted, most participants were still reluctant to concede that random assortment of plasmid molecules at division is sufficiently precise to explain the very low frequencies of plasmid-free clones which are observed even with low copy number plasmids. Several segregation or partition mechanisms were therefore proposed. In particular, the suggestion (T. Hashimoto-Gotoh, Max-Planck-Institut, Berlin) that one pair of plasmid molecules was attached to the cell membrane and partitioned accurately whereas other copies partitioned at random seemed attractive to many.

K. Nordstrom (Odense University) described *par* mutants of plasmid R1 in which copy control is maintained but a few per cent of plasmid-free clones arise in each generation. The wild type *par* function was cloned separately from *rep* functions and since the fragment failed to complement *par* mutants in *trans*, *par* may be a *cis*-acting DNA sequence essential for membrane binding and hence partition. The novel suggestion by N. Sternberg and M. Yarmolinsky (Frederick Cancer Research Center) that site-specific recombination might also play an important part in plasmid stability or

How to make a drumlin

from Ian Smalley

THOSE streamlined whale-back hills of glacial till (boulder clay) first recognised and named by M.H. Close in Ireland in 1867 still present problems. Drumlins have fascinated geomorphologists perhaps more than any other glacial landform — or so claim Whittecar and Mickelson (*J. Glaciol.* 22, 357; 1979) — their symmetry, spatial groupings, and the variability of their lengths, heights and materials have led to a variety of theories of drumlin formation. In general they are believed to have been formed by the moulding action of the ice-sheet on newly-deposited moraine.

Whittecar and Mickelson have investigated the Waukesha drumlin field which contains about 650 drumlins and related streamlined features surrounding Waukesha, Wisconsin, about 20 km west of Milwaukee. The field was formed by ice advancing out of the Lake Michigan basin approximately 14,500 years ago and it abuts the Kettle interlobate moraine to the west and is overlapped by a younger set of moraines to the east.

On the basis of the internal structures of the drumlins Whittecar and Mickelson propose a sequence of five events involved in their formation. The particular structures may be unique results formed by special situations in ice temperature regimes, groundwater flow, stratigraphy, and sediment bulk densities. Whatever the features shown, however, they often indicate that material was being moved into the drumlin from below. A growth relationship may exist that is based on the rates of material moving into the drumlin from below and rates of erosion of the drumlin form.

J. Menzies of Brock University, Ontario has also proposed a new theory of origin (*Earth Sci. Rev.* 14, 315; 1979). He suggests that the new proposal might point the way to some fruitful field research and have a unifying effect on the diverse disciplines involved in the drumlin problem. He notes that it is commonly observed that subglacial debris emerging into the proglacial zone from beneath an ice mass is slurry-like, having a very high water content. Such material would not be capable of developing into a drumlin-like obstruction at the base of an active ice

mass since it tends to be rapidly removed due to its extremely low shear strength (the result of a high pore water content) in relation to the basal shear stresses of the ice.

For such material to survive at the ice/glacier bed interface a marked change in its shear strength must occur, and the most likely process to bring about such a change is pore water dissipation. With stresses applied to the material at the interface the pore water will, if routes are available, dissipate away from points of high pressure. The rate of dissipation can be expected to vary depending on the lengths of the routes and on the size of the voids and thus on the texture and consolidation characteristics of the materials. If the water can escape from the interfacial material a relatively rigid structure might be formed which could become a drumlin. Thus the essence of the Menzies theory is that drumlin formation is controlled by variation in pore water pressure in the glacial till material.

The pore water theory bears some similarities to the dilatancy theory of drumlin formation (dilatancy is the expansion of closely packed granular masses with change in shape — in this case under ice pressure) but invokes a different mechanism to provide critical stress levels within the drumlin material. According to Menzies the dilatancy theory still holds a certain sway; he divides drumlin theorists into those who concur completely with dilatancy theory and those who do not reject the theory but who have alternative hypotheses. Crozier (*Can. Geogr.* 14, 181; 1975), studying the Peterborough drumlin field in Canada, devised a test for the dilatancy theory and claimed convincing support for it.

In a second paper (*J. Glaciol.* 22, 373; 1979) Menzies gives more details of his formation mechanism. Two mechanisms of pore-water removal are possible. First, pore water may be removed from localised patches within a mobile layer of till at the ice/glacier bed interface, thus creating nuclei of higher-strength till around which deforming till may adhere. Second, removal of water, initially from the thin water film at the base of a glacier, may result in increased pressure melting of the ice leading to till melt-out and subsequent loss of pore water from the deposited till. Menzies rather favours the first mechanism. □

Ian Smalley is with the Soil Bureau, Department of Scientific and Industrial Research, Lower Hutt, New Zealand.

partitioning was received with considerable interest. On the basis of their studies with phage P1, Sternberg and Yarmolinsky proposed that a site-specific recombinational hot-spot ensures that plasmid dimers or higher multimers created through the strong pressures for homologous recombination are likely to be

separated and therefore efficiently partitioned. In support of this C. Barton (University of Wisconsin) reported the apparent impossibility of isolating dimers of the large plasmid NR1 *in vivo*.

On a related theme J. Broome-Smith (University of Leicester) described the site-specific, *recA* independent, recombination

of ColE1 (and ColK) with a comobilisable plasmid during transfer. It was suggested that such recombination is a consequence of the normal mobilisation process, involving strand cutting at, and rejoining of, the similar transfer origins. M. Masters (Edinburgh University) also described the *recA* independent integration of plasmids bearing *ori-C* into a specific site near the origin of the bacterial chromosome and suggested that this event might also involve a site-specific endonuclease, perhaps involved in this case in initiation of chromosomal replication. Finally, R.H.

Pritchard (University of Leicester) recalled the puzzling findings, discussed at the previous workshop, that several high copy number 'synthetic' plasmids (for example, *ldv*,) and copy mutants of Clo DF13 were extremely unstable in *recA*⁺ but less so in *recA*⁻ hosts. Many participants therefore left the meeting intending to search for site-specific recombination mechanisms, perhaps hitherto recognised as *par* functions, evolved by plasmids either to reduce recombination between plasmid molecules or to resolve multimers when they occur. □

Transposable elements and chromosomal rearrangements in cancer — a possible link

from Ruth Sager

THE existence of transposable genetic elements which can insert into different chromosomal sites and thereby modulate gene expression was first described in maize by Barbara McClintock more than 25 years ago. More recently, similar elements — insertion sequences, transposons, and the mutation-inducing phage *mu* — have been found in bacteria; the molecular basis of transposition is under investigation in bacterial systems and in maize. The likelihood that the non-random chromosome changes associated with particular kinds of cancer might also be mediated by transposition events was the subject of a recent workshop entitled 'Chromosomal Transpositions as Causal Agents in Cancer'.*

Chromosome changes

Consistent chromosome changes affecting specific chromosome segments have now been revealed in banding studies of chromosomes from thousands of patients with a variety of haematological malignancies, and similar associations of chromosome aberrations with other kinds of cancer are being found with increasing frequency. J. Rowley (University of Chicago) described the chromosome gains, losses, deletions, and translocations associated with leukaemias, especially chronic myeloid and acute nonlymphocytic leukaemias; these chromosome changes arise in bone marrow cells. In the inherited form of retinoblastoma, however, patients have an abnormality of chromosome 13 in every cell. A. Mark (Central Hospital, Falk, Sweden) reported that chromosome 22 is regularly lost in human meningiomas, and that many 'double minute' chromosome fragments are present in astrocytic gliomas. In non-malignant lymphomas, short segments from other chromosomes

are inserted into a particular region of the long arm of chromosome 14 (14q+); and trisomy for chromosome 3 is also found. A similar 14q+ chromosome was also seen in B cell acute lymphocytic leukaemia.

G. Klein (Karolinska Institute, Stockholm) reported that the 14q+ chromosome seen in Burkitt's lymphoma is a reciprocal translocation with chromosome 8, and is found in tumour cells but not in the peripheral circulation. In recent studies, constant break points have been identified in both translocated chromosomes 8 and 14.

Klein proposed that the 8;14 translocation results from random changes followed by cell selection occurring in stem cells made hyperplastic by infection with Epstein-Barr virus and by additional mitogenic stimuli such as malaria. He stressed the importance of host chromosome changes in the development of virus-associated cancer. Rowley proposed that the evident specificity of break points and rearrangements in cancer-associated chromosome aberrations might result from preferential pairing of homologous DNA regions or from random rearrangements giving rise to cells with a proliferative advantage. Transposable genetic elements like those in maize have the properties that could determine the rearrangements seen in cancer progression.

Assays for tumorigenicity

Interpreting the phenotypic expression of particular chromosome changes requires cellular assays of the malignant phenotype. A. Pardee (Sidney Farber Cancer Institute, Boston) compared growth control in normal and tumour cells transformed by DNA viruses, RNA viruses and chemicals. In normal cells, growth is controlled externally by cell-cell interactions and by growth factors (proteins and small peptides) supplied by the vascular system. Tumour cells no longer respond completely to external growth controls; considerable variation is

seen from one tumour type to another in the particular responses that have been lost. Pardee proposed a unified theory of growth control in which a key protein, the 'restriction protein', regulates an on-off switch that permits or blocks entry of cells into a new round of the cell cycle. He further proposed that the rates of synthesis and degradation of this protein vary such that tumour cells generally have a higher level than normal cells but variation also exists among tumour types.

J. Ponten (Uppsala) proposed three requirements for normal cell growth: low molecular weight components; growth factors; and the solid substrate. In the presence of one of Ham's recent new media for human cells, the growth factor most clearly required by glial cells is PDGF, the platelet-derived growth factor. With respect to substrate attachment, Ponten has used circular palladium islands of controlled size to which cells attach to show that normal glial cells require a minimum of 500 square micrometres for growth regardless of growth factor availability. This result is consistent with Folkman's evidence that cell shape is of controlling importance in growth, but does not distinguish between shape and anchorage.

Ponten stressed that with glioma cell lines established from human tumours, no single phenotypic trait correlated well with tumorigenicity *per se*. He proposed that population immortality is the only reliable qualitative distinction between normal and tumour cells.

Additional cellular changes leading to metastatic potential were discussed by M. Burger (Biozentrum, Basel) who proposed that specific surface carbohydrate changes may influence the degree to which some tumour cells metastasise. Starting with the B16 mouse melanoma cell line, he used resistance to wheat germ agglutinin as a selective system to recover mutants with altered surface glycoproteins. Some variants that 'home' preferentially to the liver were also isolated for analysis on the basis of organ selectivity.

In summary, loss of growth control and immortality were taken as the best cellular evidences of transformation, and further changes were considered necessary for establishment of tumorigenicity. The relation of cell shape and anchorage to tumorigenicity remained unclear.

Chromosomal changes and malignant phenotype

Genetic analysis of cells in culture provides a means to link chromosomal changes with specific phenotypic alterations involved in malignancy. R. Sager (Sidney Farber Cancer Institute, Boston) described a new Chinese hamster cell line which is a stable diploid (22 normal

*Held in Sils Maria, Switzerland on 8-12 June, 1979 and sponsored by the International Union against Cancer and the Swiss National League against Cancer.

Ruth Sager is Professor of Cellular Genetics at the Sidney Farber Cancer Institute, Boston, Massachusetts.

chromosomes) and non-tumorigenic. Transformed and tumorigenic mutants have been recovered and genetic differences between normal and mutant cell lines have been examined by cell fusion and subsequent chromosome loss from the tetraploid back to the diploid level, during which reassortment of parental chromosomes occurs. Evidence has been found for suppression in cell hybrids of the anchorage requirement and of tumorigenicity, confirming and extending the original studies of Klein and Harris. In some reduced hybrids with 22 chromosomes, anchorage independence and tumorigenicity were re-expressed independently indicating that these traits are under separate genetic control. In addition, fusions of transformed cells with enucleated normal cells to form cybrids have shown that cytoplasmic genes are involved in tumorigenicity. The importance of examining early stages in tumorigenicity in order to identify the initial key changes that occur was stressed.

The methodology of cell hybridisation has been developed and applied by B. Ephrussi and his students to the analysis of differentiation. M. Weiss (CNRS, Gif-sur-Yvette) described the extinction of sets of phenotypic traits of rat hepatoma cells following fusion with normal or malignant mouse cells, and the subsequent re-expression of individual phenotypes coincident with loss of various chromosomes. For example, when rat hepatoma cells that produce albumin are fused with pigmented mouse melanoma cells that don't, the resulting hybrid clones are unpigmented and do not produce albumin, but they give rise to hybrid subclones that have lost chromosomes and segregate for pigmentation and albumin production. Stable mutants have been recovered from the hepatoma line that have lost some or all of their differentiated liver functions; and revertants from these mutants have regained the full set of enzymes. The reversion process is not increased by X-ray or chemical mutagens.

The extinction of sets of phenotypes in cell hybrids followed by full or partial regaining of activity, in parallel with chromosome elimination, strongly resembles suppression of tumorigenicity and its reappearance in reduced hybrids. Both systems can be interpreted in terms of regulatory signals that can turn on or turn off expression of sets of genes. As with tumorigenicity, the genetic basis of extinction and re-expression needs to be examined at the DNA level.

The relation between cancer and normal cell differentiation was stressed by E. Gateff (University of Freiburg) who described twelve unlinked *Drosophila* mutations which are larval lethals. The mutant genes interfere with development of specific tissues or cell types in a stage-specific way and thus are developmental mutations affecting rates of cell proliferation and morphogenetic

processes. In addition, neoplastic sublines have been derived from wild-type imaginal disks maintained in prolonged subculture *in vivo*. One of these lines has shown a capacity for reversal to the differentiated state, suggesting the regulation by genetic on-off switch. The wealth of descriptive detail that Gateff has produced over the past few years makes this material a valuable resource and a challenge for anyone who can develop a molecular approach to its analysis.

The possibility that the cell cycle might be regulated by one of the enzymes involved in initiation of DNA replication was raised by B. Alberts (University of California, San Francisco). He suggested that eukaryotic DNA replication might profitably be approached by looking for eukaryotic enzymes with functions analogous to the prokaryotic ones. Alberts also discussed the possibility that gene regulation might occur in some instances at the supra-DNA level of chromosome organisation.

Transposition and carcinogenesis

Of central interest to the workshop were molecular aspects of gene organisation such as the eukaryotic split genes described by P. Chambon (Strasbourg) and mechanisms of transposition in bacteria described by N. Kleckner (Harvard University, Cambridge), A. Toussaint (Brussels), and H. Saedler (Freiburg), subjects previously reviewed in *Nature*.

Despite the differences between prokaryotes and eukaryotes in gene and in chromosome organisation, unanticipated similarities have been discovered with respect to the existence and roles of transposable genomic fragments.

Saedler connected the prokaryotic and eukaryotic systems by a brief account of the discovery of chromosome transpositions in the corn plant *Zea mays* by Barbara McClintock (who unfortunately was unable to attend). McClintock discovered the existence of gene-control elements capable of inducing chromosome breaks and rearrangements and of modulating gene expression both qualitatively and quantitatively. She characterised several classes of these elements and showed that their transposition from one chromosomal location to another can be stimulated by chromosomal breakage. Whether transposition is a regular feature of their function in normal development is not known, but the transposition process is clearly under genetic control, since its time and frequency of occurrence is highly non-random and is inherited as a property of the controlling element.

The similarities in behaviour of the controlling elements in maize and the more recently discovered insertion elements and transposons of bacteria was a principal factor in the design of this workshop. In McClintock's words: "Should a genome-disturbing event appear, such as by a

cancer-inducing agent, the result could be induction of transposition of elements that can serve to control gene action. The insertion of such elements at given gene loci could initiate those genomic changes that lead to a cancerous state of the cell."

According to this view, initiation of tumorigenicity by a chemical or a virus might result from a subtle chromosome change which would precipitate subsequent chromosomal instability. Progressive development of malignancy might then depend on random karyotypic reorganisation and selection of new cell types that have lost growth control and gained other traits that facilitate growth in the given environment.

This approach to carcinogenesis may be particularly fruitful in suggesting new kinds of experiments in which the process of tumorigenesis is analysed in terms of transposition and chromosome rearrangement. Understanding the underlying karyotypic alterations that occur during tumour cell selection may provide new insights into the choice of effective anti-cancer drugs, both clinical and prophylactic. □

Pollen problems

from Peter D. Moore

WHEN interpreting the fossil pollen content of past lake sediments in terms of the contemporaneous, pollen-producing vegetation, one needs to consider the spatial origins of the pollen assemblage. Some will have arrived from the atmosphere directly, either blown from nearby plants, or carried down by drops of rainwater which effectively 'forage' for particulate matter in the atmosphere (McDonald *Science* **135**, 435; 1962). Other pollen will have been transported into the lake by the input streams.

Stream-borne pollen may itself be derived from a variety of sources. Some may have been washed from leaf and other surfaces in the catchment, where it had become impacted from the atmosphere (Chamberlain & Chadwick *Ann. appl. Biol.* **71**, 141; 1972). Some may be derived from other types of deposit, such as peat or undecomposed humus and litter from soils in the catchment. This latter source is of particular concern to the palynologist, for such pollen could be considerably older than the remainder of the pollen load in the stream and would distort the general pollen record within the lake basin. Some attempts have been made to separate this inwashed pollen component within fossil

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, King's College, London.

pollen assemblages (for example Birks *New Phytol.* **69**, 807; 1970) on the basis of the different types of damage which such grains will have suffered during transport, but generally little is known of its importance.

An experimental approach to this question has been conducted by Bonny (*J. Ecol.* **66**, 385; 1978) making use of the large tubes inserted into Blelham Tarn in the English Lake District by Lund (*Mitt. Int. Verein. Limnol.* **18**, 71; 1972). These tubes are 45 m in diameter and extend from the surface to the bottom of the lake (11–12 m) and thus effectively isolate a body of water from the remainder of the lake. Bonny established pollen traps at the lake surface and underwater, both inside and outside the tubes; traps inside the tube would be expected to receive the atmospheric pollen component, but to lack the pollen transported into the lake by streams. She found that the annual catch of pollen within the tubes was only 15% of that caught outside. This suggests that as much as 85% of the pollen submerged in the lake may have been brought there by stream transport, though this figure may be inflated by the input from marginal, eroding banks and organic deposits.

Pennington (*New Phytol.* **83**, 189; 1979) has now approached the problem from a different angle, by comparing the pollen record within the sediments of an enclosed basin lake with no inflow streams with one into which streams flow. Blelham Tarn was again taken as an example of the latter, and Whinfell Tarn, a 5.5 hectare basin also in southern Cumbria, as an example of the former. Sediment cores were extracted and the pollen stratigraphy determined using absolute techniques, covering about the past 7,000 years.

The general course of vegetation history as recorded by the pollen was similar in the two lakes. There were fewer irregularities in the enclosed basin, perhaps as a result of the smaller influence such processes as local forest clearance would have had upon erosion and pollen input. On examining absolute pollen influx into the two lakes over the past 7,000 years, however, considerable differences were apparent. Current pollen sedimentation in the surface muds of Blelham Tarn are about nine times greater than those at Whinfell Tarn. Differences between the two basins, however, have not always been so great. About 5,000 years ago, before the local deforestation of the two sites, the stream-fed lake received only twice the pollen influx of the closed basin. Following catchment deforestation this rose to about four times and the difference has been rising rather erratically ever since. This seems to be due to the erosion of organic, mor humus from the cleared forests of the Blelham catchment and the stream-borne input of this older pollen into the lakes, raising the pollen influx values by about five times. This effect is not seen in the enclosed basin. The subsequent erratic

behaviour of the influx curves at Blelham Tarn can be correlated with further soil disturbance by such processes as ploughing and cultivation.

The conclusion, therefore, is that the stream contribution to a lake's pollen content will vary with the vegetation and soils of the catchment. Where this consists of stable forest, stream-borne pollen may be less than 50% of the total pollen influx. Forest clearance or soil cultivation releases

any acidic organic matter into streams and may inflate this source of pollen to about 85% of the total input.

Such contamination with old pollen is a serious disadvantage when trying to interpret pollen data and indicates that a closed basin provides a more reliable pollen record than an open one. The alternative, of course, is to choose an ombrotrophic (rain-fed) peat site in which all pollen input is atmospheric in origin. □

Chemical recognition in biology

from James Ofengand

SPECIFICITY of reaction is the hallmark of all biological processes. In order to achieve specificity, biological molecules have learned how to recognise one another. Understanding the chemical basis for recognition has been the goal of biochemists and molecular biologists for many years, and none has been more persistent in its pursuit than Fritz Lipmann, whose 80th birthday recently was marked by an EMBO Workshop on Concepts of Chemical Recognition*. Many of the participants were former associates of Professor Lipmann, and while their interests have diverged widely over the years, the approaches taken still reflect in a characteristic way their common exposure to the view that the behaviour of biomolecules can ultimately be explained in terms of the laws of physics and chemistry.

The meeting opened with a series of talks on protein-ligand interactions. C. Walsh (Massachusetts Institute of Technology) discussed suicide substrates, structural analogues of a substrate which carry a latent reactive group capable of being exposed only by the catalytic action of the enzyme. Since these compounds are only unmasked and activated when at their target enzyme, they are well suited for *in vivo* studies as well as for consideration as chemotherapeutic agents. An example of the latter possibility is clavulanic acid (Fisher *et al. Biochemistry* **17**, 2180, 2185; 1978). This is a penicillin analogue which inactivates the β -lactamase responsible for most of the penicillin resistance encountered in pathogenic organisms. Walsh also pointed out the importance, *in vivo*, of a low partition ratio. This is the ratio of activated product which diffuses away from the active site to that which reacts with the enzyme and inactivates it. One substrate cited whose partition ratio approaches zero is the naturally occurring cyclohexadiene amino acid, gabaculine, which inactivates various GABA transaminases. The development of suicide substrates for pharmacologically important enzymes should be a rational approach to chemotherapy.

M. Perutz (MRC Laboratory of

Molecular Biology, Cambridge) drawing primarily on X-ray crystallographic data, described how haem proteins recognise their substrates. Haemoglobin and myoglobin prefer O₂ over CO because of steric interference of the side chain of HisE7 which blocks linear addition of the CO molecule but not the angular addition of O₂. This conclusion was verified using Hb Zürich, in which HisE7 is replaced by Arg. Similarly, cytochrome b₅ is an electron carrier rather than an O₂ carrier in this case because HisE7 is only 2 Å from Fe²⁺ instead of the 4 Å in haemoglobin or myoglobin so that even O₂ is sterically excluded.

Protein-protein recognition was discussed by M. Kamen (University of California, San Diego), using the recognition of cytochrome *c* by cytochrome oxidase as the paradigm. Kamen concluded that a pattern of positive surface charges produced by exposed lysine residues on one face of the cytochrome *c* molecule determines the nature of the interaction with the oxidase. Although the overall net charge on both proteins is negative, the concentration of positive charges on one surface allows electrostatic forces to pull the two molecules together, and the specific pattern of these charges determines the strength of the interaction and hence the specificity.

The complex bicyclic system for the regulation of glutamine synthetase (GS) activity was discussed by E. Stadtman (National Institutes of Health). This 12-subunit enzyme exists in an adenylylated (inactive) and non-adenylylated (active) form with each subunit able to accept one AMP residue. The interconversion requires ATP and adenylyl transferase (AT), which can both adenylylate GS with release of PP_i and de-adenylylate by phosphorolysis to yield ADP. Which of the two activities is catalysed by AT is determined by yet another protein, PII. PII is required for adenylation in addition to AT but does not induce de-adenylation. The latter reaction requires AT, P_i, and the uridylylated form of PII, PII-(UMP)₄. The interconversion of PII and PII-(UMP)₄ is accomplished by a fourth protein, uridyl transferase, which when stimulated by α -ketoglutarate, adds

*Held at the Institut National Agronomique, Grignon, France, on 18–20 July, 1979.

the UMP group of UTP to PII, and when glutamine is present hydrolyses PII-(UMP)₄ to PII. The net consequence of this elaborate system is that the α -ketoglutarate/glutamine ratio determines the activity of glutamine synthetase, although as Stadtman was careful to point out, the complexity of the system allows more than 40 different metabolites to modify the overall GS activity level. When GS is partially adenylated, 12mers with varying degrees of adenylation and varying specific activities are produced rather than a mixture of fully adenylated and nonadenylated 12mers.

The subject of the Lipmann Lecture, given by Y. Nishizuka (Kobe University, Japan), was protein kinases. After reminding the audience that Lipmann was also a pioneer in the field of protein phosphorylation (Lipmann & Levene *J. biol. Chem.* **98**, 109; 1932; Lipmann *Biochem. Zeit.* **262**, 3, 9; 1933), Nishizuka described a new class of multifunctional protein kinases discovered in his laboratory. These kinases (molecular weight 77K) are found in various mammalian tissues and require both Ca²⁺ and membranes for activity instead of cyclic AMP or cyclic GMP (Takai *et al. J. biol. Chem.* **254**, 3692; 1979; *J. Biochem.* **86**, 575; 1979). The membrane component can be replaced by certain purified phospholipids or, even more effectively, by the neutral lipid diolein. Ca²⁺ stabilises the enzyme-phospholipid complex as its removal by chelating agents dissociates the protein-lipid complex and destroys kinase activity. Interestingly, dibucaine, a local anaesthetic known to interact with phospholipids, also blocks the activity of this kinase, although it has no effect on cyclic AMP or cyclic GMP-dependent protein kinases. After proteolysis the catalytic fragment (molecular weight 51,000) is fully active without Ca²⁺ and phospholipid. The specificity of the lipid-protein interaction described here should be a fruitful area for further study in addition to the obvious possibilities for trans-membrane control implicit in a Ca²⁺ regulated membrane-associated protein kinase.

Protein-nucleic acid recognition was discussed by several speakers. G. Hartman (University of Munich) described some of his group's recent work on the recognition of promoter sequences by RNA polymerase. Although subunits from *M. luteus* and *E. coli*, two taxonomically distant species, are freely interchangeable in reconstituting overall activity, the promoter sequences recognised depend both on the source of enzyme and of the DNA, and vary from no to complete discrimination; analysis of hybrid molecules shows specificity for promoter recognition is due to the sigma subunit.

Recognition of DNA by the protein anti-

biotic, neocarzinostatin (NCS) was described by I. Goldberg (Harvard University). This antibiotic, which inhibits both bacterial and mammalian cell growth, has been used clinically as an anti-tumour agent. It is a compact protein of 10,700 molecular weight and introduces single strand breaks in DNA at T or A residues probably by splitting the deoxyribose sugar between C_{3'} and C_{4'}. A chromophoric group (molecular weight 600) in NCS may be a cofactor.

Specific recognition of satellite DNA by protein was suggested by some recent work from the laboratory of H. Zachau (University of Munich). Each long repeat unit (2,350 base pairs), is made up of four segments, each consisting of variations on the theme of a basic 23 base pair unit which is itself made from a 12 bp and a related 11 bp sequence. Analysis of the complete sequence showed that the extent of divergence as well as the frequency of base substitution at specific positions was characteristic for each segment. A high incidence of symmetry was also found, and provided the basis for Zachau's suggestion that proteins might specifically recognise certain satellite sequences.

Recent experiments on the control of replication of SV40 were described by D. Nathans (Johns Hopkins University) (Shortle & Nathans *J. molec. Biol.* **131**, 801; 1979). Using their localised mutagenesis technique (*Proc. natn. Acad. Sci. U.S.A.* **75**, 2170; 1979) they isolated mutants in the origin region of SV40 DNA, which fell into four classes corresponding to differences in rates and requirements for DNA replication, and were not correctable by complementation with wild-type gene products. Sequencing established the exact nature of the base changes and Nathans proposed that mutants in the *ori* region alter a sequence which is recognised by the T antigen, a viral protein involved in the initiation of viral DNA synthesis. This hypothesis is supported by the isolation of second site revertants some of which were mutants in the T antigen gene.

Various aspects of protein-nucleic acid recognition involved in protein biosynthesis were considered. A. Rich (Massachusetts Institute of Technology) presented some recent experiments (Wrede *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 3289; 1979) which revealed unexpected and characteristic differences between the anticodon loop structure of the initiator methionine tRNA, tRNA^{Met}_i, and the elongator tRNA, tRNA^{Met}_m in the pattern of cleavage by the single-strand specific endonuclease S1. Both prokaryotic and eukaryotic initiator tRNAs were split at two sites in the anticodon, residues 34 and 35, while elongator tRNAs of prokaryotes and eukaryotes were split at residues 33, 34, 35, and 36 (34-36 are the anticodon). The reason for this characteristic difference in splitting pattern is not known. Rich speculated that the three G-C pairs found in the anticodon stem of all initiator tRNAs, and

the only sequence feature common to all initiator tRNAs, might somehow give the anticodon loop a characteristically different structure from that of elongator tRNAs.

H. Wittmann (Max-Planck-Institut, Berlin) summarised the current state of knowledge of the *E. coli* ribosome. The primary sequence of 47 of the 53 ribosomal proteins is now known, and their shape is being studied by a variety of physical methods. A number (S1, S18, S7/12, S21, S5) are very elongated with axial ratios of 7-10:1. Several are modified by acetylation or methylation on side-chains or N-termini and three distinct N-terminal acetylases have been identified, but the purpose behind these elaborate modifications has yet to be determined. The architecture of the ribosomal particle is being studied by chemical cross-linking techniques and by immunoelectron microscopy. Two previously described crosslinks have now been precisely identified, the 8 Å link between Lys93 (S8) and Lys166 (S5), and the 0 Å crosslink between Met114 (S7) and Urd1239 (16S RNA). Ultimately, the Berlin group expects to map out most of the contact sites in the ribosome in this way. New immunoelectron microscopic results were the localisation of the 3'-end of the 16S RNA, and of the chloramphenicol binding site at the peptidyl transferase centre.

Y. Kaziro (University of Tokyo) reviewed the way that the energy of hydrolysis of GTP is used to do mechanical work in protein synthesis. Both in A site binding and in translocation, the energy comes from the participation of a protein-GTP complex which induces a change in state such that the reaction proceeds. GTP hydrolysis occurs only subsequent to the mechanical process to release its carrier protein from the ribosome by producing a conformational change in it. Kaziro suggested that this concept of work first, hydrolysis later, might be extended to other biological systems where work is done such as muscle contraction, transport, and other polymerisation reactions. Kaziro has also nearly completed sequencing of the EFTu gene, *tufA*.

Chemical recognition is being used in cancer chemotherapy to direct antimetabolites to the desired site (N. Kaplan, University of California, San Diego). 6-aminonicotinamide, a potent antimetabolite, has been coupled to polylysine to prevent it crossing the blood-brain barrier, the major cause of its toxic side effects. Polylysine coupling to methotrexate has been used to fool a methotrexate-resistant cell line in which resistance was due to a failure to transport the drug. A third example was the coupling of daunomycin, an inhibitor of DNA and RNA synthesis, to melanotropin in order to concentrate the drug specifically in the melanoma cell. This idea of chemically addressing drugs which normally have no capacity to recognise specific cell types should be a fertile ground for new approaches to chemotherapy. □

James Ofengand is a Full Member of the Roche Institute of Molecular Biology, Nutley, New Jersey.

review article

Redshifts and distances

G. Burbidge

Kitt Peak National Observatory, Tucson, Arizona

Evidence concerning the associations between QSOs and galaxies with similar redshifts, and pairs with very different redshifts is reviewed. It is concluded that both kinds of associations are likely to be real.

FOR more than 10 years there have been doubts over whether all redshifts of extragalactic objects are associated with the expansion of the Universe and can be used to indicate distances. The arguments have centred largely on the nature of the redshifts of the QSOs, although Arp¹ has given evidence which suggests that even some galaxies have non-cosmological redshifts, and Tift² has suggested that the redshifts of the galaxies in clusters, in particular, the Coma and Perseus clusters, are correlated with their nuclear magnitudes and has suggested that distances derived by the usual assumptions cannot be valid.

Because there has been no realistic theory which would account for non-cosmological redshifts, the majority feel that there is no good case for this point of view and strenuous efforts are made to disprove the existing evidence. At the same time, some evidence in favour of the cosmological interpretation has been found, and this evidence is regarded more highly.

I have previously attempted to summarise the evidence on both sides of the controversy³. Here I will concentrate on findings made largely since then on one aspect of the problem—the apparent associations between QSOs and galaxies—the first assumption being that the redshifts of the galaxies give good measures of their distances.

The investigations fall into several categories:

- (1) The programme of looking for QSOs near bright galaxies which has been carried out by Arp and his colleagues.
- (2) Searches among specified samples of pairs of QSOs.
- (3) Attempts to find associations between QSOs and faint galaxies on the assumption that the redshifts are cosmological.
- (4) Studies of the correlations between galaxy samples and QSO samples without any assumption being made as to the nature of the QSO redshifts.

Before summarising the results of these different investigations a few comments should be made. The average QSO redshift z is close to 1 and only a very small fraction of the known QSOs (~5%) have redshifts ≤ 0.2 . On the other hand, practically all of the galaxies in the published catalogues are thought to be at distances such that their cosmological redshifts are ≤ 0.2 . Thus, except in special cases in which faint galaxies are deliberately sought out and observed, in any associations that are found between QSOs and galaxies, the QSO redshifts will be much greater than that of the galaxies. A further point concerns Arp's samples. The frequent criticism is that while much is made of the peculiar objects, there must be many galaxies whose fields have been searched and in which no QSOs or anomalous redshift galaxies have been found. The implication is that the cases he describes are rare, and thus may be accidental. In fact the limits on observing time and the exacting nature of the

photographic work means that this is not true. To carry out such investigations on a large number of fields would take decades at the rate at which dark time is now available on the large telescopes. I believe that the fact that Arp finds so many of these cases suggests that the phenomenon is widespread, and that more cases are not found because no other astronomers are looking.

QSOs near to bright galaxies

Arp⁴ has listed 23 cases of QSOs found near bright galaxies. Of these 20 are NGC galaxies, two are IC galaxies and one was not catalogued. The objects are of two types: QSOs which have been found near to bright galaxies such as the 3C QSOs studied many years ago⁵, and QSOs which have been found in conjunction with fainter galaxies which are themselves thought by Arp to be associated with bright galaxies with very small redshifts. In such cases the fainter galaxies generally have redshifts much greater than those of the bright galaxies.

Since 1976 several new examples have been found by Arp and his colleagues. A list of all of the systems of galaxies and QSOs which are known has been published elsewhere⁶. Both QSOs associated with galaxies with much smaller redshifts, and fainter galaxies found around QSOs which in some cases have the same redshifts are included.

The list includes 94 large-redshift QSOs close to 55 NGC or IC galaxies with $m_v \leq 13$ and 10 fainter galaxies, several of which are in the Zwicky catalogue. All of these cases involve highly discrepant redshifts. The remainder are nearly all very faint galaxies near to QSOs with $z < 0.45$ chosen by Stockton⁷ who was looking for cases in which the redshifts agree. We shall return to them later. Among the pairs with discrepant redshifts, a few have been found accidentally, some by deliberate searches involving a limited sample of bright galaxies and radio QSOs (the 3C QSO galaxy-pairs⁵ and the PKS $\pm 4^\circ$ QSOs and Zwicky catalogue galaxies⁸) and the majority by looking in the vicinity of galaxies chosen because detailed radio studies indicated some possible peculiarity or because Arp found a peculiar fainter galaxy in the vicinity of the brighter one.

Table 1 shows all of those QSO-galaxy pairs which have separations $\leq 10'$. This table is abstracted from ref. 6. (The separation is either between a bright galaxy and a QSO or between a QSO and a companion galaxy close to a bright galaxy. These latter cases are all due to Arp.) There are 68 QSOs and 54 galaxies in Table 1. Of these, 57 are QSOs associated with NGC or IC galaxies and 11 are QSOs associated with fainter galaxies, five of which are brighter than 15.5 mag.

Given the heterogeneity of the data, the best way to evaluate the results statistically is to ask, if one looks within a range of angular distances of a chosen set of points, how many QSOs would be expected to be found by chance.

The information that we need to evaluate the result is the surface density of QSOs, and the number of cases (points) that have been examined.

It is generally agreed that the number of radio-quiet QSOs is far larger than the number of radio-emitting QSOs. The surface density of QSOs given by Sandage and Luyten⁹ and Setti and Woltjer¹⁰ down to $B = 20$ mag are 10 and 8.5 per square degree. Arp, Sulentic and de Tullio¹¹ obtain 7 ± 4 per square degree. Savage and Bolton (quoted by Wills¹²) obtain values of 0.2, 1.0 and 3.0 QSOs per square degree to $B < 18.0$, 19.0 and 20.0, respectively. Wills¹² has criticised the problem and has concluded that a value of 20 per degree down to $B = 20.5$ mag is indicated with a very considerable uncertainty. All of the surface densities mentioned above are in good agreement so that we conclude that values of 0.3 QSOs per square degree to 17 mag, 1 QSO per square degree to 18 mag, 3 QSOs per square degree to 19 mag, and 10 QSOs per square degree to 20 mag can be used. At the bright end this number is obviously conservative.

How many fields have been searched? Only Arp has looked extensively, sometimes after a peculiarity in the galaxy or companion has been recognised. For example, his study of the vicinity of NGC1073 began after Condon and Dressel¹³ reported the discovery of a compact radio source associated with a blue stellar object in the galaxy. The object was found to be a QSO, but it is the only example they found in a study of ~ 200 galaxies for compact radio sources. Thus this might well be a

chance effect. On the other hand the discovery of two more (radio-quiet) QSOs by Arp and Sulentic¹⁴ may be highly significant.

The observing time required to obtain adequate photographic material compared with the limited dark time available for this work leads me to conclude, with Arp, that a maximum of 100 fields have been studied by him and his colleagues. Some of the other objects have been found by accident, and we shall assume at first that the total number of fields that have ever been looked at is 200.

Let $\Gamma(<m)$ be the sky density of QSOs brighter than m . Then the expected number lying within an angular separation θ of arbitrary points is

$$\langle n \rangle = 2.4 \times 10^{-7} N \Gamma(<m) \theta^2$$

where N is the number of tries, Γ is expressed in $(\text{arc deg})^{-2}$ and θ in arc s.

Table 2 gives the number of QSOs expected by chance, as a function of separation from the galaxy and as a function of magnitude of the QSOs, compared with the total number found (N_0). The number that has been found by Arp is designated N_{0A} .

Table 2 shows that the number found within 1, 2 and 3 arc min is far greater than the number expected by chance and thus the result is highly significant.

It may well be argued that the number of tries used to calculate $\langle n \rangle$ can only reasonably be compared with N_{0A} , the number of pairs discovered by Arp and that comparison of N_0 with $\langle n \rangle$ is unfair, because N_0 includes QSOs found close to galaxies in other surveys and more than 200 tries are involved. If we take this more conservative point of view the result has

Table 1 All QSOs known to lie within 10 arc min of bright galaxies or their companions

QSO	Galaxy	Separation (arc s)	m_v (QSO)	z (QSO)	QSO	Galaxy	Separation (arc s)	m_v (QSO)	z (QSO)
UM228	NGC78B	516	17.0	0.102	UB1	NGC3384	265	19.4	1.111
UM245	NGC120	240	18.0	(1.46)	UB13	NGC3384	149	20.6	(0.497)
UM247	NGC132	300	17.0	2.35	4C61.20	NGC3407	173	16.3	0.422
BSO1	NGC157	119	19.0	0.756	1107+036	Anon	20	18.9	0.96
	(companion)				1108+289	NGC3561	66	19.0	2.192
UM268	NGC227	456	18.0	(1.66)	BSO1	NGC3726	100	18.4	1.13
PKS0038-019	NGC227	67	18.0	1.690		(companion)			
	(companion)				W1	NGC3992	280	20.3	1.75
PKS0038-020	NGC227	~ 400	18.0	1.178		(companion)			
	(companion)				1205-008	Anon	9.4	18.6	1.002
PKS0106+013	ZW0106.1+0123	192	18.39	2.107	3CR268.4	NGC4138	174	18.4	1.400
PKS0112-017	NGC448	600	18.0	1.365	MK205	NGC4319	42	14.5	0.070
0121+108,	Anon	40	18.0	0.510	UB1	NGC4395	145	18.7	1.265
MC2						(companion)			
UB1	NGC622	71	18.5	0.91	B6	NGC4395	329	18.4	1.038
BSO1	NGC622	73	20.2	1.46		(companion)			
PHL1226	IC1746	54	17.5	0.404	3CR270.1	NGC4395	303	18.6	1.519
UB2	NGC772	352	19.4	2.61		(companion)			
0219+428	3C66B	390	15.2	(0.44)	UB1	NGC4395	370	19.2	0.77
PKS0225-014	NGC936	600	18.0	2.037	Weedman 8	NGC4550	44	17.2	0.728
BSO1	NGC1073	104	19.8	1.945	3CR275.1	NGC4651	210	19.0	0.557
BSO2	NGC1073	117	18.8	0.599	1253+104 MC2	Anon	90	18.0	0.824
RSO	NGC1073	84	20.0	1.411	UB1	NGC5055	315	18.3	0.910
UB1	NGC1087	170	19.1	2.147		(companion)			
4C-02.15	NGC1298	228	19.5	2.092	BSO1	NGC5296	55	19.3	0.963
0844+319	NGC2402	30	18.87	1.834	UB2	NGC5660	483	17.3	0.205
0846+51	NGC2681	12	19.5	1.860	UB1	NGC5669	70	17.7	0.766
	(companion)					(companion)			
U2	NGC2859	66	(19.7)	2.25	BSO1	NGC5682	95	19.2	1.94
	(companion A)				1441+522	3C303	20	19.97	1.570
U3	NGC2859	73	(20.3)	1.46	3CR309.1	NGC5832	372	16.8	0.905
	(companion B)				BSO1	NGC5866	436	18.1	0.706
U1	NGC2859	60	(19.2)	0.23	UB1	NGC5981	107	19.0	2.132
	(companion C)				1548+114B	Anon	12	19.0	1.901
0925+301	B20925+301	480	21.0	2.02	W1	NGC6503	324	16.5	(0.76)
HOAG 1	M82	382	20.0	2.0492	2020-370	Anon	21	17.5	1.048
HOAG 2	M82	512	21.0	(2.1)	BSO1	IC1417	76	17.8	0.73
HOAG 3	M82	585	21.0	2.0440	3CR455	NGC7413	24	19.7	0.543
3C232	NGC3067	114	15.8	0.534	UB1	NGC7465	128	19.2	1.66
UB2	NGC3338	251	19.7	2.14	UB1	NGC7592	415	18.7	1.410
	(companion)				UB1	NGC7714-15	120	18.0	1.871
PKS1021-006	ZW1022.0-0036	126	18.22	2.547	UB2	NGC7714-15	480	19.0	2.193

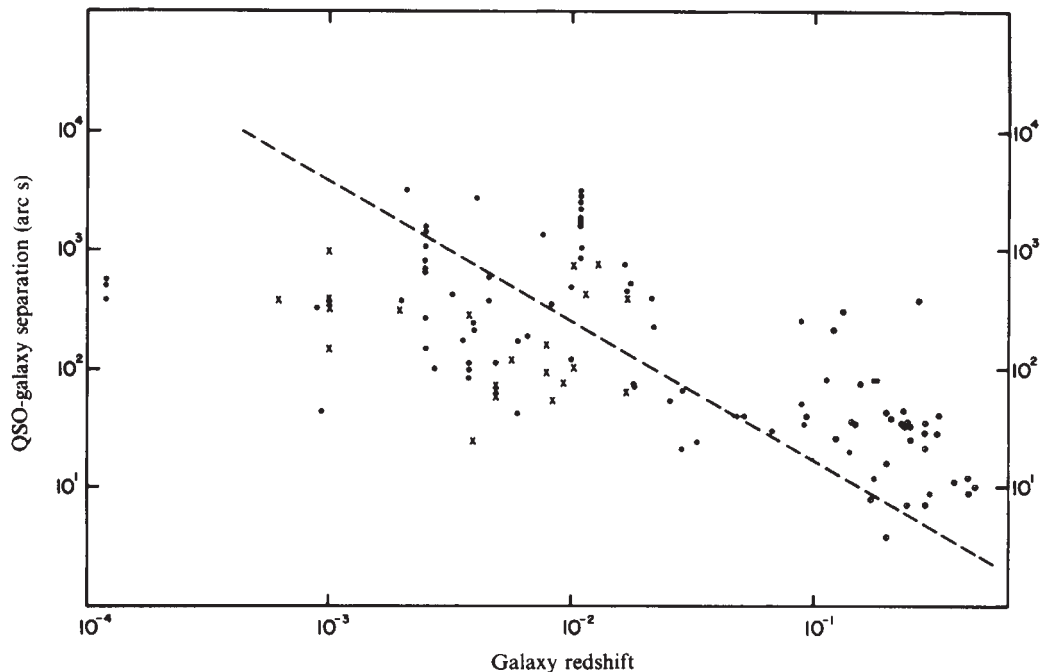


Fig. 1 A plot of $\log z_{\text{galaxy}}$ against $\log \theta$ for all apparent QSO-galaxy associations listed by Hewitt and Burbidge⁶. The low redshift QSOs marked with a separate symbol are the QSOs close to galaxies investigated by Stockton³². The line is the least-squares solution which has the form $\log \theta = -1.17 \log z_{\text{galaxy}} + \text{constant}$. The range of slope at the 99% confidence level lies between -0.93 and -1.56 . ●, Main galaxy; ×, redshift (main) versus θ (companion); ○, low redshift QSO.

obviously less statistical weight. Note that N_0 also includes the five 3CR QSOs found near to bright galaxies^{5,8,15} which were shown to be highly significant⁵ as a separate sample. These associations were found by studying galaxies in the vicinity of about 50 3CR QSOs. Thus only 50 tries were involved and our total of 200 tries assumed above can reasonably include that sample. Also included are the associations between Parkes QSOs and Zwicky catalogue galaxies¹⁸. Here about 90 QSOs were involved. The other sample which is included in N_0 are the cases recently found by Weedman¹⁶ who has looked for associations between 259 QSOs found by objective prism techniques and 62 bright galaxies in the same area of the sky. He found that the apparent associations lay within one standard deviation of what one would expect by chance. The numbers that are seen as one goes to larger distances ($\sim > 5'$) are comparable with or even less than the numbers expected by chance. This may be because Arp has not searched as hard for QSOs further out. Alternatively it may be that the mean surface density of QSOs that we have used is too high. If the latter is correct, the result based on the number of close QSOs is enhanced.

What about NGC1073 where three QSOs are found to lie within $2'$ of the centre of that galaxy. Let P_m be the probability that a single galaxy has m QSOs within θ arc by chance. Then

$$P_m(\mu) = \frac{\mu^m e^{-\mu}}{m!}, \quad \mu = \frac{\langle n \rangle}{N}$$

From Table 2 we find that there are 34 galaxy-QSO pairs with separation $\leq 3'$. Because in the case of NGC1073 it was the first radio QSO which drew attention to the galaxy, we evaluate the probability that two QSOs are found within $3'$ of a point which already has one QSO within $3'$. Then $P = 0.0028$, and the probability that any one of the 34 galaxies with one QSO within $3'$ will have two more there by accident is 0.095.

Alternatively we might ask what is the probability that any of the $\sim 2,000$ galaxies in the revised Shapely Ames catalogue has three QSOs lying within $3'$ of its nucleus by chance. If $m = 3$, $P_m(\mu) = 7.22 \times 10^{-5}$. Thus the total probability is 0.144. This assumes that the vicinities of all of the bright galaxies have been searched for QSOs but only a very small number, perhaps 100, have been studied in this way.

Note also that the redshifts of the three QSOs associated with NGC1073 are 1.945, 0.599 and 1.411 (Table 1). This might be another accident, but it is striking that these values are exactly at the peaks of the redshift distribution discussed over many years (see ref. 17).

Many of the QSOs found near to bright galaxies by Arp and his colleagues are not close enough to be statistically significant if we estimate the probabilities in the normal way. But in many of these cases alignment of pairs of QSOs across the nucleus of a galaxy and the similarity of the redshifts of the QSOs makes physical association very likely. Alignments of this type include the QSO and two other compact radio sources associated with the radio galaxy B20925+30 (refs 4, 18), the QSOs around NGC3384 (ref. 19), QSOs around NGC2639 (ref. 20) and the pairs of aligned triple QSOs without a bright galaxy present found by Arp and Hazard²¹. All of the systems of this type have been listed elsewhere⁶.

Given that there is some statistical evidence of association of galaxies and QSOs for close separations, and that alignments and similarity of redshifts of several pairs of QSOs around bright galaxies suggest physical association, we can look for further evidence by making a plot of the redshifts (proportional to distance) of the galaxies against angular separations of the QSOs from the galaxies. When this was originally done for the 3CR QSOs⁸ it was found that the slope of the $\log z_{\text{galaxy}} - \log \theta$ relationship was very close to -1 , suggesting that the QSOs were at approximately a constant projected distance from the galaxies. Figure 1 shows a similar plot for all of the galaxy-QSO systems listed by Hewitt and Burbidge. Included are 94 QSOs associated with 65 galaxies. Obviously some of these may be accidental and because in many cases several QSOs at different distances are associated with the same galaxy, we expect considerable scatter in the relationship. However, a least squares solution gives a best fit line for $\log z$ given $\log \theta$ of the form

$$\log \theta = -1.17 \log z + \text{constant}$$

The range of slope for this line at the 99% confidence level is between -0.93 and -1.56 . The correlation coefficient associated with the $(\log \theta, \log z)$ pairs is $r = 0.68$. This correlation is significant.

Pairs of QSOs

Stockton²² found the first close pair of QSOs separated by 35", with very different redshifts. He calculated *a posteriori* that the probability that this was due to chance was 6×10^{-3} . The next pair that was found was 1548 + 114a and b, with a separation of 5" (ref. 23), *a posteriori* probability calculation gave the probability of a chance superposition $\sim 10^{-2}$. Bolton *et al.*²⁴ carried out a systematic search of 100 fields centred on flat spectrum radio-emitting QSOs from the Parkes 2,700 MHz survey. They found eight candidate QSOs within 1.7' of the radio-emitting QSOs, and have confirmed that there are at least four and probably five pairs of which at least four and probably five pairs have different redshifts. As the plate limit in these studies was $B = 19-19.5$ mag, the number of QSOs per square degree is about three, and thus the number of pairs expected by chance is 0.76. Since four or five pairs are found, Bolton *et al.* consider that the probability that this is due to chance lies in the range of $6 \times 10^{-3}-10^{-4}$. Wills and Wills¹² then did a similar survey in the north, searching within 2' of 100 QSOs identified in the NRAO 5,000 MHz surveys. They found two pairs and one is expected by chance. Thus from the two fields together, seven pairs have been found against the 1.76 expected by chance.

Several new close pairs or triplets have recently been found. They include a pair with separation 5.7" with precisely the same redshift²⁵ which must be a physical pair, or a single object which appears double due to an intervening massive object²⁵, and the triple system of QSOs near M82²⁶ with separations of 75" and 125". Each QSO has a redshift close to 2.04.

Perhaps we are seeing several phenomena in the close groups, including chance effects, physical pairs with the same redshift, physical pairs with different redshifts, and physical pairs or triplets with similar large redshifts ejected from nearby bright galaxies. Only further observations will clarify this situation.

Associations between QSOs and galaxies at the same redshift

The first successful attempt to find a galaxy at the same redshift as a QSO was made by Gunn²⁷, and Robinson and Wampler²⁸ who found one or two normal galaxies lying near the QSO PKS2251 + 11. One at least has a redshift very close to 0.323 which is the redshift of PKS2251 + 11. Gunn made statistical arguments *a posteriori*, concluding that there was a very low probability that this association was accidental and this result was heralded as evidence that the redshifts of the QSOs are of cosmological origin. Several other similar cases were found²⁹⁻³². However, there was no reliable way of estimating the statistical significance of these associations³³.

Recently Stockton⁷ has carried out a much more significant study. He chose all of the known QSOs (27) with redshifts ≤ 0.45 , $m_v < 19.2 + 5 \log z$ and at declinations $-15 < \delta < +55$ and attempted to obtain spectra of all galaxies visible on the red Palomar Sky Survey plates lying within 45" of the QSOs. Of the 29 galaxies he identified, he obtained spectra of 25 and has reported that 13 galaxies in eight fields (out of 27) have redshifts, cz , within $1,000 \text{ km s}^{-1}$ of the redshift reported for the QSO. He has estimated that the probability that this could occur by chance

is less than 1.5×10^{-6} . Stockton³⁰ also has found a galaxy at the same redshift as 3C273, although it lies further away than 45".

These results provide strong evidence that some QSOs at least lie at cosmological distances.

General studies of correlations between faint galaxies samples and QSOs

Several attempts have been made to look for correlations between the positions of QSOs and faint galaxies contained in the Abell catalogue and the Lick catalogues. The maximum redshift for any of these galaxies is $z \sim 0.2$.

Bogart and Wagoner³⁴ found marginal but not convincing evidence for an apparent correlation between the positions of QSOs with large redshifts ($z \geq 0.3$) and Abell clusters ($z \leq 0.2$). Recently Roberts *et al.*³⁵ investigated the positions of all QSOs with measured redshifts (498 objects) in the area of the sky containing the Abell catalogue of rich clusters. It was found that there is no statistically significant evidence that either low- or high-redshift QSOs lie preferentially close to Abell clusters. This is in contrast to the highly significant evidence that 3CR radio galaxies lie preferentially close to Abell clusters.

This result probably supersedes the early result of Bogart and Wagoner because much larger samples are involved. It does not contradict the work of Stockton as the galaxies he found close to his sample of low redshift QSOs are rarely, if ever, in rich clusters.

Seldner and Peebles³⁶ have found statistically significant evidence of a correlation between the position of 382 QSOs at $|b| \geq 40^\circ$ $\delta > -23^\circ$ and the Lick counts of galaxies, $m \leq 19$. They find on the average 1.45 ± 0.39 more galaxies within 15' of a QSO than would be expected if the QSOs were placed at random. There seems to be no trend with QSO redshift or apparent magnitude and the effect is not apparently due to a few QSOs at the same redshifts or the galaxies of the type found by Stockton. Again this result does not conflict with the absence of correlations with Abell clusters as the majority of the galaxies in the Lick counts lie in small groups and more extended distributions.

From these results and those of Stockton we conclude: first, that QSOs are not present in rich clusters; second that QSOs are sometimes physically associated with small groups of galaxies. In some cases they may have the same redshifts, with no significant non-cosmological component, while in others they may be a large non-cosmological component.

Conclusion

Based on the evidence discussed here some QSOs apparently have redshifts which are completely cosmological in origin, but others have redshifts of largely non-cosmological origin. It is hard to deny either the evidence of Stockton, or the evidence of QSOs found close to bright galaxies or their companions.

The results therefore imply either that there are two distinct classes of QSOs with very different physical redshift mechanisms operating, or that the Doppler effect is still always responsible. In this latter situation both cosmological Doppler shifts and those due to high speed ejection of QSOs from galaxies must be involved.

Table 2 Comparison between the numbers of QSOs found near to bright galaxies (N_0 or N_{0A} and the number expected by chance ($\langle n \rangle$)

Apparent magnitude	$\theta \leq 60''$		$61'' \leq \theta \leq 120''$		$121'' \leq \theta \leq 180''$		$181'' \leq \theta \leq 300''$	
	N_0	$\langle n \rangle$	N_{0A}	$\langle n \rangle$	N_0	$\langle n \rangle$	N_{0A}	$\langle n \rangle$
≤ 17	1	0.05	0	0.16	0	0.24	1	0.82
≤ 18	5	0.17	0	0.52	4	0.86	1	2.8
≤ 19	9	0.52	1	1.55	9	2.4	2	8.2
≤ 20	12	1.73	3	5.2	12	8.6	4	27.6

Apparent magnitude	$301'' \leq \theta \leq 600''$		$\theta \leq 180''$		$\theta \leq 600''$	
	N_0	$\langle n \rangle$	N_{0A}	$\langle n \rangle$	N_0	$\langle n \rangle$
≤ 17	4	3.9	1	0.46	1	5.2
≤ 18	9	13.0	2	1.56	5	17.3
≤ 19	15	39.1	8	4.62	12	51.9
≤ 20	18	129.6	10	15.5	19	173

I thank H. C. Arp for discussions and for data before publication. J. V. Narlikar looked again at the correlation between galaxy redshifts and angular separations, I thank him for permission to reproduce Fig. 1. I also thank A. Hewitt for help in the preparation of the paper. The Kitt Peak National Observatory is operated by the Association of Universities for Research Inc., under contract with the NSF.

1. Arp, H. *Science* **151**, 1214–1216 (1966).
2. Tift, W. G. *Astrophys. J.* **175**, 613–625 (1972); **179**, 29–44 (1973); **181**, 305–326 (1973).
3. Burbidge, G. R. *Nature phys. Sci.* **246**, 17–25 (1973).
4. Arp, H. *Colloq. Int. No. 263*, 377–398 (CNRS, Paris, 1977).
5. Burbidge, E. M., Burbidge, G. R., Solmon, P. M. & Strittmatter, P. A. *Astrophys. J.* **170**, 233–240 (1971).
6. Hewitt, A. & Burbidge, G. *Astrophys. J. Suppl.* (submitted).
7. Stockton, A. N. *Astrophys. J.* **223**, 747–757 (1978).
8. Burbidge, G. R., O'Dell, S. L. & Strittmatter, P. A. *Astrophys. J.* **175**, 601–611 (1972).
9. Sandage, A. & Luyten, W. J. *Astrophys. J.* **155**, 913–918 (1969).
10. Setti, G. & Woltjer, L. *Ann. N.Y. Acad. Sci.* **224**, 8–30 (1973).
11. Arp, H., Sulentic, J. W. & de Tullio, G. *Astrophys. J.* **229**, 489 (1979).
12. Wills, D. *Phys. Scr.* **17**, 333–337 (1978).
13. Condon, J. J. & Dressell, L. L. *Astrophys. J.* **221**, 456–467 (1978).
14. Arp, H. & Sulentic, J. W. *Astrophys. J.* **229**, 496 (1979).
15. Arp, H. C., Burbidge, E. M., Mackay, C. D. & Strittmatter, P. A. *Astrophys. J. Lett.* **171**, L41–L43 (1972).
16. Weedman, D. W. Preprint (1979).
17. Burbidge, G. *Phys. Scr.* **17**, 237 (1978).
18. Ekers, R. D., Fanti, R., Lari, C. & Ulrich, M.-H. *Nature* **258**, 584–586 (1975).
19. Arp, H. *Ann. N.Y. Acad. Sci.* (in the press).
20. Arp, H. Preprint (1979).
21. Arp, H. & Hazard, C. Preprint (1979).
22. Stockton, A. N. *Nature phys. Sci.* **238**, 37 (1972).
23. Wampler, E. J., Baldwin, J. A., Burke, W. L., Robinson, L. B. & Hazard, C. *Nature* **246**, 203–205 (1973).
24. Bolton, J. G., Peterson, B. A., Wills, B. J. & Wills, D. *Astrophys. J. Lett.* **210**, L1–L3 (1976).
25. Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381 (1979).
26. Burbidge, E. M., Hoag, A. A., Junkkarinen, V. & Smith, H. E. (in preparation).
27. Gunn, J. C. *Astrophys. J. Lett.* **164**, L113–L118 (1971).
28. Robinson, L. B. & Wampler, E. J. *Astrophys. J. Lett.* **171**, L83–L86 (1972).
29. Stockton, A. N. *Nature phys. Sci.* **246**, 25 (1973).
30. Oemler, A., Gunn, J. C. & Oke, J. B. *Astrophys. J. Lett.* **176**, L47–L50 (1972).
31. Stockton, A. N. *Nature* **250**, 308 (1974).
32. Stockton, A. N. *Astrophys. J. Lett.* **205**, L113–L116 (1976).
33. Burbidge, G. R. & O'Dell, S. L. *Astrophys. J. Lett.* **182**, L47–L51 (1973).
34. Bogart, R. S. & Wagoner, R. V. *Astrophys. J.* **181**, 609–618 (1973).
35. Roberts, D. H., O'Dell, S. L. & Burbidge, G. R. *Astrophys. J.* **216**, 227–236 (1977).
36. Seldner, M. & Peebles, P. J. E. *Astrophys. J.* **227**, 30–36 (1979).

articles

A theory of terrestrial catastrophism

W. M. Napier & S. V. M. Clube

Royal Observatory, Blackford Hill, Edinburgh, UK

The time sequence of terrestrial catastrophes and other Solar System phenomena is apparently stochastic but with an underlying galactic modulation. A mechanism is proposed involving the capture of planetesimals as the Sun passes through spiral arms.

THE radial separation of galactic spiral arms and the epicyclic frequency of the solar orbit imply quasiperiodic intervals between crossings of a few 10^7 yr. Should galactic evolution involve the formation of successive generations of spiral arms (see ref. 1), the intervals will be modulated by an additional quasiperiodicity of a few 10^8 yr. If spiral arms were to contain planetesimals, each Solar System passage would lead to the capture and formation of a temporary 'Oort cloud' which would be rapidly dissipated by planetary encounters and hyperbolic ejection. We shall argue that the cratering flux record, the current distribution of comets and the existence of many dynamically short-lived bodies in the Solar System, are each best explained as consequences of this single mechanism. With this viewpoint, comets and most asteroids and satellites are temporary survivors of one or more capture episodes. The impacting planetesimals can be shown also to cause, *inter alia*, water erosion on Mars, intermittent biostratigraphically recognisable events in Earth history (including mass extinctions) and ice ages. The latest rapidly declining series of events follows our passage through the Orion spiral arm. As the Gould Belt complex, at a mean distance of ~ 250 pc, is the nearest appendage of the Orion arm to the Sun (see ref. 2), and has a relative velocity of ~ 25 km s $^{-1}$, this capture episode ended ≤ 10 Myr ago. The most recent spell of ice ages, within the past ~ 20 Myr was preceded by a more temperate period³ lasting ~ 40 Myr. Furthermore, assuming the Orion arm has a typical mass content ($\sim 10^6$ – 10^7 $M_\odot$ kpc $^{-1}$), the calculated remaining flux of low mass

planetesimals agrees with the observed 'short period' (Apollo) asteroids and an impact rate inferred from very recent terrestrial events.

Cratering flux history

The record of impact cratering on the Moon and inner planets⁴ shows the planetesimal flux to have declined steeply for the first 2,000 Myr of the Solar System's existence, and then to have attained a constant value over the past 2,500–3,500 Myr. Superimposed on this trend are quasiperiodic bombardments, of uncertain magnitude and duration, separated by time intervals characteristically a few 10^8 yr. A similar periodicity is evident in the deposition of micrometeorites on lunar soil⁵.

Dynamical studies⁶ indicate that the steeply declining phase of the cratering history can be understood in terms of a sweeping-up, or hyperbolic ejection, of primordial debris by the planets. Rapid depletion of the inner Solar System debris occurs so that after $\sim 10^8$ yr, bodies entering the inner Solar System have semi-major axes primarily $\sim 10^5$ AU. The existence of these high initial flux rates is clear evidence that the Solar System has been immersed in a cloud of planetesimals in the past.

Two important features of the cratering history remain to be explained: the brief periods of enhanced bombardment; and the remarkable constancy, in the mean, of the planetesimal flux taken over the past $\sim 3,000$ Myr.

There are numerous perturbing sources in the interstellar medium. Thus the Solar System has passed through ~ 15 nebulae of radii a few pc and $n_H > 10^3$ in the past 4,500 Myr, and experienced ~ 60 close encounters (say with impact parameter $p < 8$ pc) over the same period⁷. Such nebulae have masses $M \sim 10^4 M_\odot$, and encounters with them would result in drastic perturbations of the planetesimal cloud and consequent flooding of the inner Solar System. This could account for the bombardment episodes. However, the very efficacy of the process raises

Table 1 Percentage of outer Oort cloud planetesimals ejected through close passage of a massive nebula*

M ($10^4 M_\odot$)	p (pc)	V (km s^{-1})	%
1	5	30	36
2	5	30	50
1	2	10	90

* Based on three-body numerical integrations (program by M. Staniucha).

difficulties for the survival of a cloud virtually intact for half the lifetime of the Solar System, as is implied by the constant flux.

Numerical trials have been run in each of which 2,500 massless bodies bound to the Sun were perturbed by a passing nebula idealised as a point mass of prescribed (M, p, V), where V represents the flyby velocity. The orbits were randomly orientated and the semi-major axes a lay in the range $10^4 \leq a \leq 2 \times 10^5$ AU, with eccentricities e in the range $0 \leq e < 1$; the test particles were therefore on the outer fringes of the hypothetical Oort cloud. The results of the three-body integrations are shown in Table 1. Clearly the outer part of the Oort cloud, from which long-period comets are supposed to derive, would be very efficiently cleared by passing dense nebulae, and it does not seem likely that the comet-producing regions of the cloud could survive ~ 30 such encounters, let alone yield a constant cratering flux for the past $\sim 3,000$ Myr.

Even if we disregard these difficulties, the hypothetical Oort cloud as currently envisaged is unlikely to be the source of impacting planetesimals. Thus, its mass is supposed to be a few $m_\oplus$ based on a comet mass of $\sim 10^{17}$ g and the likely number of comets^{8,9}. The crater size distribution¹⁰ reveals a mass function $\propto m^{-1.6}$ with a typical upper limit of $m > 10^{20}$ g (for example the Tycho crater was formed ~ 400 Myr ago¹¹ by a mass of $\sim 10^{19}$ g). The bulk of the mass of the system must therefore be concentrated in the few largest bodies whereas the 'characteristic mass' of a comet is inferred from the more common bodies at the low-mass end of the spectrum. Note that some observers¹² give independent reasons for supposing a maximum cometary mass $\sim 10^{20}$ g. It follows that the mean mass of a comet should be increased by a factor ~ 40 giving $80 m_\oplus$, say, for the present Oort cloud. The process which fed comets into the Oort cloud would have thrown the bulk of the material into hyperbolic orbits¹³. Thus $80 m_\oplus$ represents a total ejection of $> 1,300 m_\oplus$ and probably $> 2,700 m_\oplus$. These amounts are much too large to be dynamically possible for the current Solar System configuration.

These results indicate that the cratering flux must comprise two components: one a decaying part due to the primordial swarm of bodies bound to the Sun, the other a constant part which has become dominant over the past $\sim 3,000$ Myr as the first decayed. The main belt asteroids fail by a factor of ~ 10 to supply the impacting planetesimals¹⁴ and would in any case be a decaying source. The only other possible location for the constant flux source is interstellar space. In this view, the long period comets have to be identified as the most easily visible component of the captured planetesimals. (The idea that comets are interstellar in origin was first suggested by Laplace¹⁵.) The fact that there are few certain hyperbolic orbits amongst the long period comets indicates the Sun is not currently immersed in a significant interstellar source¹⁶. The observed flux is thus the residue of captures effected in the recent past.

The hypothetical Oort cloud should be thermalised by weak stellar encounters¹⁷ on a time scale ~ 1 Myr. The perihelia q should thus be uniformly distributed both over the sky and in distance from the Sun. The perihelion distribution of long period comets in fact increases as $q^{1/2}$ in the inner Solar System¹⁸, lies in a preferred plane $\sim 20^\circ$ from the Milky Way, and concentrates $\sim 7^\circ$ from the solar apex¹⁹. There is thus good evidence amongst the comets favouring a capture mechanism related to the Sun's motion through the local galactic environment.

Probably the dominant capture mechanism is through entry of planetesimals into Jupiter's sphere of influence: thus Radzievskii and Tomanov²⁰ have shown that Jupiter can capture $\sim 10^5$ planetesimals of mean mass 10^{15} g during the passage, lasting 0.1 Myr, through a cloud whose planetesimal component has a smoothed-out mean density 10^{-20} – 10^{-22} g cm⁻³. For mean mass 4×10^{17} g the required densities are 4×10^{-18} – 4×10^{-20} g cm⁻³. These densities are high but not unreasonable; and if planetesimals are widely distributed through spiral arms rather than concentrated in dense clumps, the requisite densities are correspondingly relaxed.

Yabushita²¹ has shown that the energy distribution of comets in the Solar System evolves to a form which, after 3–6 Myr, is almost independent of the initial distribution and is concentrated in the range $2 \times 10^4 < a < \infty$ AU. To be consistent with observation, capture must have taken place between 3.6 and ~ 9 Myr ago. Within the errors, this agrees satisfactorily with the emergence ≤ 10 Myr ago of the Sun from Gould's Belt.

We conclude that the constant average flux of planetesimals during the past $\sim 3,000$ Myr is ultimately of interstellar origin, and that the fluctuations in the bombardment rate reflect passage through successive spiral arms, the most recent of which is responsible for the currently observed declining flux of comets.

Some implications of interstellar planetesimals

If our hypothesis is correct, then planetesimals, possibly up to at least lunar size, must occur in abundance in spiral arms. Note that a depletion of C, N and O observed in the interstellar medium has been ascribed to the existence of interstellar comets²² and has certain implications.

First theories of the origin of the Solar System involving accumulation of blobs rather than, say, condensation of a solar nebula, become more plausible. Star formation as a process of aggregating small masses rather than fragmenting large ones is then probable. Although such ideas^{23,24} have never found favour, the alternative of Jeans collapse is proving unsatisfactory²⁵. The hypothesis is *a priori* consistent with star formation occurring predominantly in spiral arms. The planetesimal mass function indicated by the crater size distribution is developed in a fragmenting-accreting regime²⁶ and must be consistent with the presence of dust in spiral arms. Cold 'globules' have been observed²⁷, and these provide evidence that dusty protostellar objects of the kind implicit in this scheme of star formation do actually exist. These objects are also notable sources of CO emission²⁸, and not only is CO emission generally very concentrated in the galactic plane where spiral arms lie²⁹, but it is commonly detected in regions of intense star formation³⁰. We emphasise therefore that although the existence of planetesimals is not generally considered to be an essential requirement for any theory of how stars are formed, it is, nevertheless, consistent with observation.

A deeper problem arises when we consider the origin of the planetesimal material. The interstellar accretion processes stu-

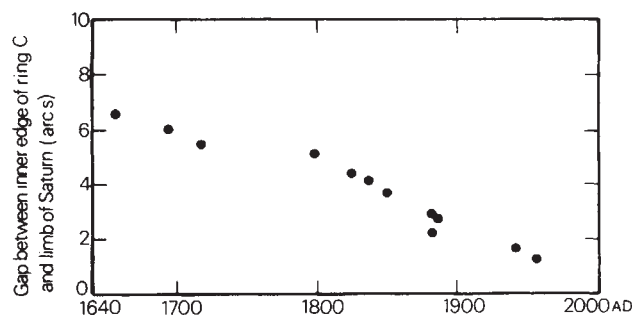


Fig 1 Gap between inner edge of ring C and limb of Saturn, measured in arc s and normalised to a mean opposition. Secular evolution of the whole ring system on a time scale a few 10^4 yr seems probable.

Table 2 Short-lived or recent Solar System phenomena

Object	Phenomenon	Time scale (yr)	Probability*	Comments	Refs
Chiron	Hyperbolic ejection from Solar System	$\sim 5 \times 10^4$	$\sim 10^{-5}$		53, 54
Hidalgo	Hyperbolic ejection from Solar System	$\sim 10^6$	$\sim 2 \times 10^{-4}$		55
Triton	Spiralling in towards planet	10^7 – 10^8	10^{-3} – 10^{-2}	Largest satellite ($\sim 3,700$ km diameter); retrograde orbit, indicative of capture	36, 37
Phobos	Spiralling in towards planet	$\sim 6 \times 10^7$	$\sim 10^{-2}$		38
Rings of Saturn	Secular evolution	$\sim 10^2$ – 10^4	$\sim 10^{-6}$	See Fig. 1. Ring C is spreading inwards at ~ 1.6 arc s per yr	39–41
Mimas/Tethys resonance	Evolving	$\sim 2 \times 10^8$	$\sim 4 \times 10^{-2}$	2:3 resonance in period, disturbed or formed a few 10^8 yr BP	42
Enceladus/Dione resonance	Evolving	$\sim 4.4 \times 10^8$	$\sim 10^{-1}$		43
Outer retrograde jovian satellites	Temporary capture	?	?	Orbits are retrograde, indicative of capture, and are unstable in Hill's sense; asteroids are an inadequate replenishing source	44–48†

* Probability is time scale divided by 4.5×10^9 yr.

† A. E. Roy, personal communication.

died so far³¹ give growth times of ~ 100 Myr to attain particle diameters $\sim 10^{-5}$ cm. Some forced diffusion mechanism (see refs 32, 33) of remarkable efficacy is required if, say, density waves were the prime agency in forming spiral arms. But if spiral arms were the result of some exotic process, say involving recurrent (galactic) nuclear explosions, then one might look for the origin of planetesimals in that same exotic process. Seyfert events may be the result of such explosions at 10^8 yr intervals³⁴, and these can be understood in terms of very massive stars ($\geq 10^8 M_\odot$) evolving through and beyond a nuclear burning phase³⁵. On this hypothesis therefore, planetesimals are suddenly cooled fragments of central highly evolved, very massive stars¹.

Recent events and anomalous phenomena in the Solar System

One might expect to find manifestations of interstellar capture events among the small body population of the Solar System. Indeed we propose that in addition to the comets, many asteroids and satellites form a temporary population in a state of relatively rapid (10–100 Myr) dynamic evolution, being the survivors of one or more capture episodes.

Table 2 lists some short-lived Solar System phenomena. Qualitatively, these find a unified explanation in terms of a temporary capture scenario; taken cumulatively, they seem otherwise inexplicable. For example the probability, on a conservative view, that we happen to simultaneously observe the Triton, Phobos and Chiron events is $\sim 2 \times 10^{-10}$. Singer⁴⁹ has argued that all the regular satellites of the major planets were captured after the formation of the planets themselves. Currently, direct orbits predominate because retrograde orbits decay more rapidly. The slope of the satellite mass function is similar to that of asteroids, in agreement with a common origin hypothesis. The largest satellite masses are also close to that which comes within the capture radius of Jupiter during spiral arm passages.

A fundamental problem of the asteroid system is that it has much more internal kinetic energy than can be fed into it by jovian perturbations. Jupiter can induce eccentricities of at most a few hundredths in the asteroid system, substantially less than those observed (which are up to ~ 0.40)⁵⁰. Another manifestation of this problem⁵¹ results from application of Tisserand's criterion to Pallas: it has been shown that this large asteroid (and others) cannot have been perturbed by Jupiter from the invariable plane into its high-inclination orbit—yet

conventionally, accretion could have occurred only in the quiescent conditions supposed to pertain in the invariable plane⁵².

These problems do not arise on the capture hypothesis, according to which the asteroid zone represents a relatively stable region of phase space, and the asteroids are, in part or whole, interstellar planetesimals which have accumulated during successive bombardment episodes. One therefore expects the asteroids to occupy every available stable niche of phase space, as is observed. The observed chemical differentiation is then a matter for conjecture: the outer, carbonaceous asteroids may simply not yet have lost their surface coating; or we may be observing the products of successive capture events from interstellar regions with locally different properties.

Short-period comets are regarded as deriving from intermediate and long-period comets through jovian perturbations⁵⁶. As they have lifetimes of only a few centuries then one expects dynamic equilibrium to be established in time scales of that order. A recent study¹⁸ of the number density of long-period comets in the Solar System indicates that this is constant throughout the inner regions, and is $\sim 0.09 \text{ AU}^{-3}$. With this figure only 1 comet per 11 yr is expected to enter the 'capture region' around Jupiter. The probability that a comet, entering the capture region, will actually be captured⁵⁷, is $\sim 7 \times 10^{-3}$. One therefore expects an interval $\sim 1,600$ yr between captures of long-period comets into short-period orbits. Since the lifetime⁵⁸ before disintegration of a short-period comet is ~ 500 yr, then there is an expectation on the 'steady state' model of observing ~ 0.3 short-period comets. The number actually observed ~ 80 .

The explanation may be that a moderately large comet (say diameter 20 km) disintegrated due to tidal forces while in the process of being captured into a short-period orbit. The maximum fragment size would be determined by the tensile strength of ice and the closeness of passage to Jupiter: fragmentation into say $\sim 10^3$ pieces of diameters ≤ 2 km is feasible.

This discussion implies that the intermittent capture hypothesis provides a single, simple explanation of diverse phenomena which otherwise require a variety of *ad hoc* theories.

Planetary impacts

The terrestrial cratering history of the Phanerozoic¹⁰ indicates that the smoothed interval between formation of craters of diameter $D \geq 20$ km, is $\delta t \approx 5.7 \times 10^5 (D/20)^2$ yr, with D

measured in kilometres. The kinetic energy E of impact is related to D by $E = 2.4 D^{3.4}$, valid for $D \geq 14$ km. (It is convenient here to measure energy in megatons equivalent of TNT: one megaton $\equiv 1$ MT $= 4.2 \times 10^{22}$ erg.) Given E , a density of 3.3 g cm^{-3} and a r.m.s. impact speed of 24.6 km s^{-1} , the projectile diameter d follows. Hence $(\delta t, d, E)$ can be calculated in terms of D and are listed in Table 3.

The actual frequency of impact is bunched into episodes of spiral arm immersion; hence from Table 3, the Earth is expected to suffer one impact of energy at least 1.6×10^8 MT during each passage of the Sun through a spiral arm; an impact with $E > 5 \times 10^8$ MT during a passage is expected with probability ~ 0.5 ; and an impact of energy $\geq 10^9$ MT is expected sometime during the Phanerozoic. These intervals should be compared, for example, with the mean duration of a biostratigraphic period (~ 50 Myr). These large impacts must inevitably have global consequences, directly through the effects of blast (see ref. 59), and indirectly through the injection of dust⁶⁰ and NO into the stratosphere.

For land impacts, the smallest overpressure likely to be experienced anywhere on Earth is $\sim 0.25 \text{ kg wt cm}^{-2}$ when, for example, $E \sim 5 \times 10^8$ MT. Peak overpressures of this magnitude completely destroy forests⁶¹. As large land vertebrates of small strength/weight ratio would be preferentially killed⁶², such a blast wave could have been the prime cause of the Cretaceous-Tertiary extinction^{63,60}. The survival of animal life associated with underground habitats and the rapid regeneration of vegetation from roots and seeds would be consistent with this hypothesis. An impact of this energy would also excavate $\sim 10^{18}$ g of material much of which would reach the stratosphere as a fine dust and block out sunlight, with catastrophic effects on food chains⁶³. If the settling time were ≥ 10 yr, the build up of ice could be ~ 0.3 km in depth and continental glaciation might be initiated⁶⁴. Analogy with nuclear explosions suggests that $\geq 10^{15}$ g of NO would be injected into the stratosphere⁶⁶, thereby destroying the ozone layer. The biological effects of the UV deluge to be expected as the sky clears are likely to be drastic⁶⁷. Ocean impacts, causing waves up to 8 km amplitude, will generally be more common than land impacts, and through simultaneous flooding of the land and obscuration of the upper atmosphere are likely to be an extremely efficient cause of sudden ice ages. An impact in the Palaeozoic Pacific at the close of the Frasnian explains⁶⁸ all the major aspects of an event which included the sudden extinction of shallow sea life.

Impacts of the sort expected at ~ 50 Myr intervals therefore produce an abundance of deleterious effects, the consequences of which are catastrophic extinctions. The situation is likely to be more complex for lesser impacts but the gradual extinction of taxa might be initiated indirectly: thus the widespread brachiopod extinctions near the Ordovician-Silurian boundary are correlated with general glaciation and withdrawal of the sea from continental platforms⁶⁹. A quasi-periodicity in the numbers of taxa is clearly evident⁷⁰ and is ~ 50 Myr. Geophysical consequences are more conjectural, but the correlation between extinction events, field reversals and rotation period⁷¹ may have impacts as a common cause. It is possible that frequent ice ages so alter the mass distribution on plates as to raise or lower land levels⁷², precipitate orogenic activity at their perimeters, and even provide a driving force for tectonics. Urey⁷³ suggested that impacting comets might induce the termination of geological periods and pointed out a correlation with tektite

Table 4 Ages of geological periods compared with tektite ages and ages of all known Phanerozoic craters > 50 km diameter

Biostratigraphic interval	Age at beginning (Myr)	Tektite/crater age (Myr)	Tektite/crater
Pleistocene	1	0.77 ± 0.10 0.88 ± 0.93	Australites Ivory Coast
Pliocene	13	14.7 ± 0.7	Moldavites
Miocene	25	28.6 ± 2	Libyan desert glass
Oligocene	36	34.7 ± 2 38 ± 9	Bediasites Popigai, 100 km, USSR
Eocene	58	57	Kara, 50 km USSR
Palaeocene	63	?	?
Cretaceous	135	?	?
Jurassic	181	183 ± 3	Puchezh-Katunki, 80 km, USSR
Triassic	230	210 ± 4	Manicougan, 70 km, Canada
Permian	280	?	?
Carboniferous	345	365 ± 7	Siljan, 52 km, Sweden
Devonian	405	?	?

ages. Table 4 lists the ages of the known tektites, and all known craters of diameter > 50 km, against the ages of the geological periods. As the record must be grossly incomplete, it is surprising that, within the uncertainties, tentative correlations can be made, agreement becoming poorer the more remote the epoch.

Comets may comprise $\sim 40\%$ water ice by mass⁷⁴, and this may hold for a proportion of the larger planetesimals: the low densities of the saturnian satellites, for example, may indicate an icy constitution. The impact hypothesis therefore offers a natural explanation for the extensive water erosion on Mars, which is known to have arisen from one or more catastrophic flooding events⁷⁵.

A prediction of the hypothesis is that anomalous isotopic and abundance anomalies should occur, because of global dust settling, at biostratigraphic boundaries. Thus Pt, Ir and Os are typically 10^3 times more abundant in extra-terrestrial than crustal material. W. Alvarez *et al.* (personal communication) have found that, in a 1-cm clay layer marking the Cretaceous-Tertiary boundary at Gubbio, Italy, the Ir abundance jumps by a factor ~ 20 , decaying upwards with a half height of 8 cm. The absence of ^{244}Pu , at ~ 0.1 of the level predicted by nucleosynthetic theories, seems to preclude a supernova source.

Discussion

Clearly the picture presented here needs clarifying. In particular the implications of large amounts of planetesimal material in spiral arms, and the terrestrial consequences of large impacts, need to be explored further. Nevertheless, starting from (and justifying) the interstellar capture hypothesis, we have been able to unify a wide range of galactic, Solar System and terrestrial phenomena; and we have identified the most recent capture event as having ended ≤ 10 Myr ago with the emergence of the Solar System from Gould's Belt.

Our theory enables the flux of planetesimals on the Earth at present, corresponding essentially to the short period core of the Apollo asteroids (ref. 14) to be calculated. This flux results in one 10^{3+1} MT event every $\sim 10^4$ yr and one of these may well have been responsible for initiating the short ice age from which Earth emerged some 10^4 yr ago. (There is a 20% probability that the event lies in the upper half of the stated range.) For comparison, the Krakatoa volcano⁷⁶ had an energy of 50 MT and the Tunguska meteorite⁷⁷ ~ 10 MT: these caused $\sim 10\%$ and $\sim 1\%$ drops in the solar constant respectively over ~ 1 yr. A stochastic distribution of similar ice ages over the preceding $\sim 10^7$ yr is implied, although with additional modulation by the Milankovitch process (see ref. 78). With the usual planetesimal mass function, the flux also corresponds to one 10^{2+1} MT event per $\sim 1,000$ yr, each with possibly detectable consequences.

Table 3 Interval δt between impacts with planetesimals of diameter $\geq d$, with impact energies $\geq E$ and producing craters of diameter $\geq D$

D (km)	δt (yr)	E (MT)	d (km)
80	9.2×10^6	7.1×10^6	3.9
100	1.4×10^7	1.5×10^7	5.0
200	5.8×10^7	1.6×10^8	10.9
500	3.6×10^8	3.6×10^9	30.8

Thus, we might expect these to produce a temporary hemispheric climatic cooling and, following calculations by Brown and Hughes⁷⁹, a correlated enhancement of ¹⁴C. Discrepancies in ¹⁴C chronology during the past several thousand years are known to occur and may coincide with 'mini' ice ages⁶⁵: a common cause for such effects seems probable. Planetary impacts have not been previously suspected in this context, and yet their impact frequency is apparently such that historical records are a potential source of data for testing our hypothesis.

Received 7 June; accepted 2 October 1979.

1. Clube, S. V. M. *Vistas Astr.* **22**, 77–118 (1978).
2. Clube, S. V. M. *Mon. Not. R. astr. Soc.* **137**, 189–203 (1967).
3. Öpik, E. J. *The Planet Earth* (ed. Bates, D. R.) 164–194 (Pergamon, New York, 1976).
4. Hartmann, W. K. *Scient. Am.* **236**, 84–99 (1977).
5. Lindsay, J. F. & Srnka, L. J. *Nature* **257**, 776–777 (1975).
6. Wetherill, G. *Proc. lun. Sci. Conf.* **6**, 1539–1561 (1975).
7. Talbot, R. J. & Newman, M. J. *Astrophys. J. Suppl.* **34**, 295–308 (1977).
8. Öpik, E. J. *IAU Symp.* 523–574 (Liege, 1965).
9. Safronov, V. S. in *Comets, Asteroids, Meteorites* (ed. Delsemme, A. H.) 483–484 (The University of Toledo, 1977).
10. Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230–242 (1979).
11. Hartmann, W. K. *Icarus* **13**, 299–301 (1970).
12. Russell, H. N., Dugan, R. S. & Stewart, J. Q. *Astronomy*, 446 (Ginn, New York, 1945).
13. Fernandez, J. A. *Icarus* **34**, 173–181 (1978).
14. Wetherill, G. W. *Icarus* **37**, 96–112 (1979).
15. Laplace, P. *Sur les Comètes, Connaissance des Temps* (1806).
16. Sekanina, Z. *Icarus* **27**, 123–133 (1976).
17. Oort, J. H. *Bull. Astr. Inst. Neth.* **11**, 91–110 (1950).
18. Kresak, L. & Pittich, E. M. *Bull. astr. Inst. Csl.* **26**, 299–309 (1978).
19. Oja, H. *Astrophys. J.* **43**, 317–319 (1975).
20. Radzievskii, V. V. & Tomanov, V. P. *Astr. Zh.* **54**, 388–397 (1977).
21. Yabushita, S. *Mon. Not. R. astr. Soc.* **187**, 445–462 (1979).
22. Greenberg, J. M. *Astrophys. J.* **189**, L81–85 (1974).
23. Krat, V. A. *Problems of Cosmology* Vol. 1, 34 (Academy of Science, Moscow, 1952).
24. Horedt, G. P. *Astrophys. Space Sci.* **45**, 353–367 (1976).
25. Larson, R. B. *Mon. Not. R. astr. Soc.* **184**, 69–85 (1978).
26. Napier, W. M. & Dodd, R. J. *Mon. Not. R. astr. Soc.* **166**, 469–489 (1974).
27. Bok, B. J. *The Moon and the Planets* **19**, 153–155 (1978).
28. Martin, R. N. & Bennett, A. M. *Astrophys. J. Suppl.* **36**, 1–51 (1978).
29. Burton, W. B. & Gordon, M. A. *Astrophys. J.* **207**, L189–193 (1976).
30. Thaddeus, P. *IAU Symp.* **75**, 37–54 (1977).
31. Wickramasinghe, N. C., Kahn, F. D. & Mezger, P. G. *Interstellar Matter*, 266 (Geneva Observatory, 1972).
32. Wesson, P. S. *Astrophys. Space Sci.* **23**, 227–255 (1972).
33. McCrea, W. H. *Observatory* **95**, 239–255 (1975).
34. Bailey, M. E. & Clube, S. V. M. *Nature* **275**, 278–282 (1978).
35. Ozerov, L. M. & Uslov, V. V. *Astrophys. Space Sci.* **13**, 3–35 (1971); **25**, 149–194 (1978).
36. McCord, T. B. *Astr. J.* **71**, 585–590 (1966).
37. Trafton, L. *Astrophys. J.* **193**, 477–480 (1974).
38. Sinclair, A. T. *Vistas Astr.* **22**, 133–140 (1978).
39. Struve, O. *Mon. Not. R. astr. Soc.* **13**, 22 (1852).
40. Vsekhsvyatsy, S. K. *Astr. Zh.* **39**, 290–302 (1962).
41. Cook, A. F. & Franklin, F. A. *Smithson. Contr. Astrophys.* **2**, 377–383 (1958).
42. Allan, R. R. *Astr. J.* **74**, 497–506 (1969).
43. Sinclair, A. T. *Mon. Not. R. astr. Soc.* **160**, 169–187 (1972).
44. Szebehely, V. *Celest. Mech.* **18**, 383–389 (1978).
45. Byl, J. & Ovenden, M. W. *Mon. Not. R. astr. Soc.* **173**, 579–584 (1975).
46. Öpik, E. J. *Interplanetary Encounters* (Elsevier, Amsterdam, 1975).
47. Heppenheimer, T. A. *Icarus* **24**, 172–180 (1975).
48. Heppenheimer, T. A. & Porco, C. *Icarus* **30**, 385–401 (1977).
49. Singer, S. F. in *Comets, Asteroids, Meteorites* (ed. A. H. Delsemme) 109–117 (The University of Toledo, 1977).
50. Heppenheimer, T. A. *The Origin of the Solar System* (ed. S. Dermott) (Wiley, New York, 1977).
51. Whipple, F. L., Lecar, M. & Franklin, F. A. in *On the Origin of the Solar System* (ed. Reeves, H.) 312–319 (CNRS, Paris 1972).
52. Safronov, V. S. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*, 25–27 (Israel Program for Scientific Translation, Jerusalem, 1972).
53. Hughes, D. W. *Nature* **270**, 385–386 (1977).
54. Oikawa, S. & Everhart, E. *Astr. J.* **84**, 134–139 (1979).
55. Öpik, E. J. *Interplanetary Encounters*, 81 (Elsevier, Amsterdam, 1975).
56. Everhart, E. *Astr. J.* **78**, 329–337 (1973).
57. Everhart, E. *Astrophys. Lett.* **10**, 131–135 (1972).
58. Oort, J. H. *The Solar System* (eds Middlehurst, B. M. & Kuiper, G. P.) Vol. 4, 665–673 (University of Chicago, 1963).
59. Croft, S. K. *Impact and Explosion Cratering* (eds Roddy, D. J. et al.) 1279–1296 (Pergamon, New York, 1977).
60. Rehfsuss, D. E., Michael, D., Anselmo, J. C. & Kincheloe, N. K. *Impact and Explosion Cratering* (eds Roddy, D. J. et al.) 1123–1132 (Pergamon, New York, 1977).
61. Zotkin, I. T. & Tsikulin, M. A. *Sov. Phys. Dokl.* **11**, 183 (1966).
62. Glasstone, S. & Dolan, P. J. *The Effects of Nuclear Weapons* 3rd edn, 559 (U.S. Government Printing Office, 1977).
63. Russell, D. A. *Cretaceous–Tertiary Extinctions and Possible Terrestrial and Extraterrestrial Causes* (eds The K-Tec group) 11–23 (National Museums of Canada, 1976).
64. Hoyle, F. & Wickramasinghe, C. *Astrophys. Space Sci.* **53**, 523–526 (1978).
65. Lamb, H. H. *Climate: Present, Past and Future* Vol 2, Part II (Methuen, London, 1977).
66. Viewing, D. R. J. & Horswell, C. J. *J. Br. Interplanet. Soc.* **31**, 209–216 (1978).
67. Ruderman, M. A. *Science* **184**, 1079–1081 (1974).
68. McLaren, D. J. *J. Palaeontol.* **44**, 801–815 (1970).
69. Boucot, A. J. *Evolution and Extinction Rate Controls*, 122–129 (Elsevier, Amsterdam, 1975).
70. Cutbill, J. L. & Funnell, B. M. in *The Fossil Record* (eds Harland, W. B. et al.) 791–820 (Geol. Soc., London, 1967).
71. Whyte, M. A. *Nature* **267**, 679–682 (1977).
72. Belousov, V. V. *Basic Problems in Geotectonics* (McGraw-Hill, New York, 1962).
73. Urey, H. C. *Nature* **242**, 32–33 (1973).
74. Delsemme, A. H. in *Comets, Asteroids, Meteorites* (Delsemme, A. H.) 3–13 (The University of Toledo, 1977).
75. Komar, P. D. *Icarus* **37**, 156–181 (1979).
76. Ben-Menahen, A. *Phys. Earth planet. Inter.* **11**, 1–35 (1975).
77. Gaskell, T. F. *Physics of the Earth*, 165 (Smeets, Weert, 1970).
78. Ruddiman, W. F. & McIntyre, A. *Science* **204**, 171–173 (1979).
79. Brown, J. C. & Hughes, D. W. *Nature* **268**, 512–514 (1977).
80. Russell, D. A. A. *Rev. Earth planet. Sci.* **7**, 163–182 (1979).

Cooperative effects in simulated water

P. Barnes, J. L. Finney, J. D. Nicholas & J. E. Quinn

Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1, UK

Computer simulation calculations on liquid water, aqueous solutions and biological interfaces have been made almost universally assuming pair-additive interactions. Experimental and theoretical evidence implies this is a serious oversimplification. The development and testing of a different approach is reviewed in which many-body forces are specifically taken into account through molecular polarisation. Solution and interfacial properties of aqueous systems are particularly sensitive to such forces.

DURING the past decade there have been significant advances in our understanding of the molecular structure and dynamics of the liquid state. This has been due largely to the development of computer simulation techniques such as Monte Carlo and molecular dynamics methods^{1–3} by which we avoid the problems remaining in analytical theories of simple liquids⁴.

The success of these methods applied to the simple liquids of the theoretician (such as liquid argon) has encouraged applications to more complex liquids. In particular, the universal relevance of water in chemistry and biology, coupled with the

fascination of its 'abnormal properties', have made aqueous systems a popular target for simulation methods^{5–9}. But if results are to have more than a hypothetical relevance to the real liquid, the potential we use must model adequately the behaviour of real water molecules. We discuss some of the difficulties facing realistic simulations of 'water-like water' which are unavoidable if solution effects and the operation of biological systems are to be understood properly. We pay special attention to the nature of water–water hydrogen bonding and argue that the long suspected cooperative nature of this interaction¹⁰ must be taken into account.

Pair additivity

A variety of potentials has been used to model the water–water hydrogen bonding interaction, such as the four-point-charge ST2 (ref. 11) potential which was derived from an earlier effective pair potential (BNS)¹², which was itself fitted to the measured second virial coefficient of steam. Other potentials in use include central-force models (CF) such as that of Lemberg and Stillinger⁹ which treats the hydrogens and oxygens as separate charged units, and one developed by Clementi by

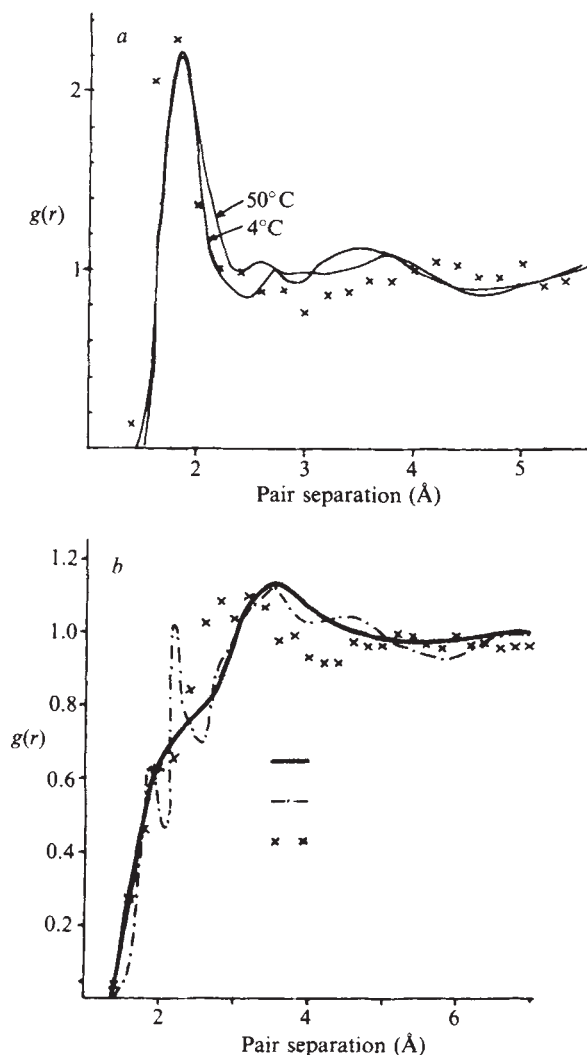


Fig. 1 *a*, Pair distribution functions, $g(r)$, obtained from X-ray scattering data⁴⁶ on water at 4°C and 50°C (—) and from simulated PE-water (×). The starting configurations were generated from an assembly of 216 molecules (previously equilibrated under the ST2 potential) with 40,000 Monte Carlo steps using a cubic box length of 18.62 Å, normal periodic boundaries and a cut off of 7–7.5 Å. *b*, A comparison of the neutron pair distribution function of D_2O at 11°C (—) with simulation predictions using the central-force (CF) model of Stillinger and Rahman⁵¹ (---) and the polarisable electropole (PE) model (×). The PE model reproduces the low- r shoulder and is notably free of the strong peaking incorrectly predicted by the central force model (which indicates the over-tetrahedrality of the latter potential).

fitting an analytical form to quantum mechanical calculations on the water dimer⁷.

All these potentials are pair-additive: the energy of a given assembly is expressed as a sum of interaction energies between all pairs of molecules in the assembly. In other words, the interaction between an isolated pair of molecules is assumed to be unaffected by the intrusion of a third molecule. This is a universal assumption when dealing with simple liquids, where it is an excellent approximation. For water it has been argued that non-pair-additive effects, although accepted to be greater than in simple liquids, can be either neglected or somehow averaged¹².

However, it has been stressed¹³ that none of these pair potentials can account for both gas and condensed phase properties of water, and recent neutron diffraction data^{14,15} on D_2O demonstrates that current four-point charge and central-force models predict an overstructured liquid. Moreover, the known failure of these simple models to account for dielectric proper-

ties is a serious drawback to understanding solvent effects in real aqueous chemical or biological systems. The following observations underline the inadequacy of pair-additive potentials: (1) Isolated water molecules are poor hydrogen donors and acceptors; in the gas phase, water is less acidic than toluene, less basic than formic acid¹⁶. This strongly implies that its celebrated acid/base properties must relate to its bulk phase, rather than to the isolated molecule. (2) Quantum mechanical calculations on, for example, water trimers and tetramers, suggest that hydrogen bonding in these clusters is stronger than in the dimer by 20–30%^{17,18}. (3) The dipole moment of the isolated water molecule is 1.85 D (ref. 19), yet the average value for hexagonal ice I is between 2.6 and 3.0 D (ref. 20). Such a large dipole moment enhancement is highly relevant to dielectric and solution properties.

One might argue that an effective pair potential might be constructed to reproduce such condensed phase effects; in view of the known extent of the non-pair additivity, such a procedure is not only artificial but is incapable of reconciling gas and liquid phase properties and therefore most unlikely to be able to predict localised effects such as at interfaces. This evidence forces us to take the cooperative effects into account specifically, and to examine their consequences.

The polarisable electropole water model

The electronic charge distribution of an isolated water molecule is represented by a multipole expansion as described by Buckingham²³ using the known dipole and quadrupole moments plus a spherically symmetric Lennard-Jones function. In developing a polarisable electropole (PE) water model^{21,22}, an electropole formalism (see Table 1 legend) was chosen in preference to a four-point charge model, (1) to avoid the need for artificial switching devices¹¹ necessary to prevent charge overlaps, (2) to ensure the correct sign and magnitude of the quadrupole moments (ST2, for example, yields a quadrupole moment of the wrong sign), and (3) to minimise the need to fit adjustable parameters. The non-pair additivity is handled using a classical polarisability tensor.

The use of this classical device requires explanation, considering that any significant covalency would result in appreciable charge transfer between two hydrogen-bonded molecules. However, as the size of the basis set used in *ab initio* quantum mechanical calculations on the water dimer is increased, the amount of charge transfer predicted falls to a degree where electronic cloud distortion into the hydrogen bond region is high but actual intermolecular electron transfer is small²⁸. A polarisation algorithm can reproduce the major effects of this charge distortion.

In performing polarisation calculations, the resultant electric (dipole+quadrupole) field acting on a given molecule will induce an electron cloud distortion quantified by the product of the field and the polarisability. This polarisation gives rise to corrections to the previously calculated fields and hence to a further change in the calculated polarisation, and so on. The total polarisation is thus determined in an iterative fashion

Table 1 Comparison of second virial coefficients

Temperature (°C)	Stock- mayer*	Model	CI†	BNS‡	ST2†	PE§	Experimental (cm ³ mol ⁻¹)		
							34	56	33
100	-450	-827	-466	-1,023	-507	-450	-460	-507	
200	-205	-289	-190	-355	-231	-197	-210	-214	
300	-122	-141	-107	-177	-125	-112	-115	-113	
400	-80	-78	-65	-104	-78	-72	-72	-72	

*Ref. 32. †Ref. 57. ‡Ref. 12.

§The present model uses an isolated dipole moment of 1.855 D (ref. 19), absolute (not derived²³) quadrupole moments ($Q_{xx} = -7.482$, $Q_{yy} = -4.180$, $Q_{zz} = -5.939$ D Å, axes as defined by Glaeser and Coulson²⁶) from quantum mechanical calculations²⁴ and an isotropic polarisability ($\alpha_{xx} = \alpha_{yy} = \alpha_{zz} = 1.444$ Å³) obtained from extrapolated²⁵ refractive index data. The Lennard-Jones parameters ($\sigma = 3.024$ Å, $\epsilon = 0.764$ kcal mol⁻¹) were chosen to fit the equilibrium dimer²⁷ properties.

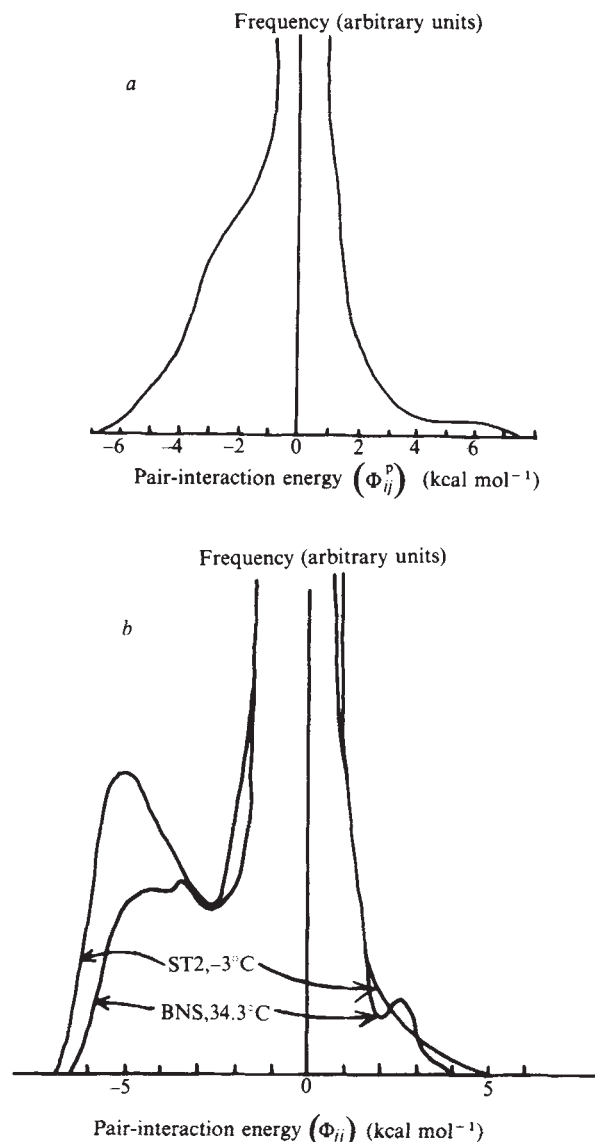


Fig. 2 The population distributions of interaction energies between all pairs of molecules in simulated configurations obtained using: *a*, the polarisable electropole; and *b*, four-point charge models. The four-point charge data (BNS⁶ at 34.3 °C, ST2 (ref. 11) at -3 °C) show a definite low energy peak or shoulder at -3 to -6 kcal mol⁻¹ which has been identified with 'hydrogen-bonded' pairs. In contrast the PE model indicates merely a broadening consistent with a picture of continuously-varying hydrogen bond energies rather than one in which bonds are discretely formed or broken.

(essentially a self-consistent field approach) and 2-5 iterations are typically necessary to achieve a constancy of better than 1% in the final charge distributions. Once this is completed, the energy is calculated by summing the pair interactions of electric fields with the resulting instantaneous dipole moments and field gradients with quadrupoles. The relevant equations are given elsewhere⁴².

The model has the following advantages: (1) The non-pair-additivity calculations are exact. No effective averaging is used. Consequently this potential can be used for gas, liquid and crystal calculations as well as interfacial and mixed aqueous systems. (2) No adjustable electrical parameters are used, in contrast to other models^{7,9,11,12,29} where parameters are adjusted to one or two sets of data. (3) Artificial devices such as switching functions¹¹ are not necessary. (4) Problems of generalising the switching functions to polarisable water molecules do not exist; electrostatic polarisation catastrophes are not

experienced in practice except in cases of pathological configurations. (5) The properties of the isolated water molecule are by definition correct. In particular the quadrupole components, in contrast to four-point-charge models, are correct in both sign and magnitude.

Tests of the model

Before using any potential in expensive liquid-phase simulations it is only sensible to test it against known properties which are easily quantified^{13,30}. If such tests give poor agreement with experiment, then the potential is clearly inadequate. One cannot overstate the need to be critical at this stage. Thus, we have preferred to gain confidence in the predictive ability of a potential function over a wide range of densities before testing it on expensive liquid, solution and biological interface calculations. The following is a selection of the tests made so far.

The second virial coefficient of steam $B(T)$: this two-molecule property, exploring the interactions of an isolated pair of molecules, can be written for a temperature T as

$$B(T) = -\frac{1}{16\pi^2} \int (\exp - \phi_{ij}/kT - 1) d\tau$$

Here, ϕ_{ij} is the interaction energy of the pair, and k the Boltzmann's constant. The integration, ideally performed over all possible configurations, raises sampling problems which are discussed elsewhere³¹. Table 1 shows $B(T)$ for several potentials. The BNS and Stockmayer potentials were obtained by fitting adjustable parameters to $B(T)$ whereas the ST2, CI and our PE model values are predictions made without such adjustments. Of the latter set, only the PE model gives a satisfactory fit, especially at low temperature.

The equilibrium water dimer: the polarisable electropole (PE) predictions are in excellent agreement with the equilibrium configuration and force constants from both quantum mechanical calculations^{7,17,18,28,36} and experimental measurements²⁷ (see Table 2).

Small water clusters: as agreement between different quantum mechanical calculations is poor, we discuss only the relative stabilities of the different cluster geometries, and the overall trends in their respective non-pair-additivity energies rather than the less reliable absolute values. Referring to Table 3, the most stable arrangements are the sequential (OH—OH—OH—) form with a distinct preference for the more closed chain varieties, and (except for the trimer) for cyclic clusters over chains. The PE model reproduces the same relative trends in equilibrium geometry and non-pair-additivity. The hydrogen-bond energy increases from -6.09 to -10.35 kcal mol⁻¹ as the cluster chain grows from two to five molecules; the enhancement of some 70% is more than a mere perturbation. Water is clearly cooperative, and the PE model reproduces structural and energetic consequences well within the variations found between different quantum mechanical calculations.

Ices: theoretical estimates of the average dipole moment $\bar{\mu}$ in ice Ih suggest a value of 2.6 D (ref. 37)—an enhancement of 40% over the isolated molecule value—while a recent experiment proposes values between 2.45 and 3.0 D (ref. 20). The PE model predicts an average value of 2.88 D, with a half-width spread of about 0.3 D. (The PE model tests were performed on an enlarged unit cell of 432 ice molecules using normal periodic boundary conditions and a cut-off distance of 11 Å. Canonical Pauling arrangements for hydrogen positions were obtained using a disordering 'loop generation' technique described elsewhere^{11,22}.)

Compared to the experimental value of -11.3 kcal mol⁻¹ for the internal energy of ice Ih (refs 38, 39), the PE model gives a value of -11.8 kcal mol⁻¹, of which -17.1 kcal mol⁻¹ comes from electrostatic contributions. Although this agreement is encouraging, we regard it as largely coincidental in view of the unoptimised repulsive core parameters. The model is also consistent with the non-existence of a ferro-electric phase of ice

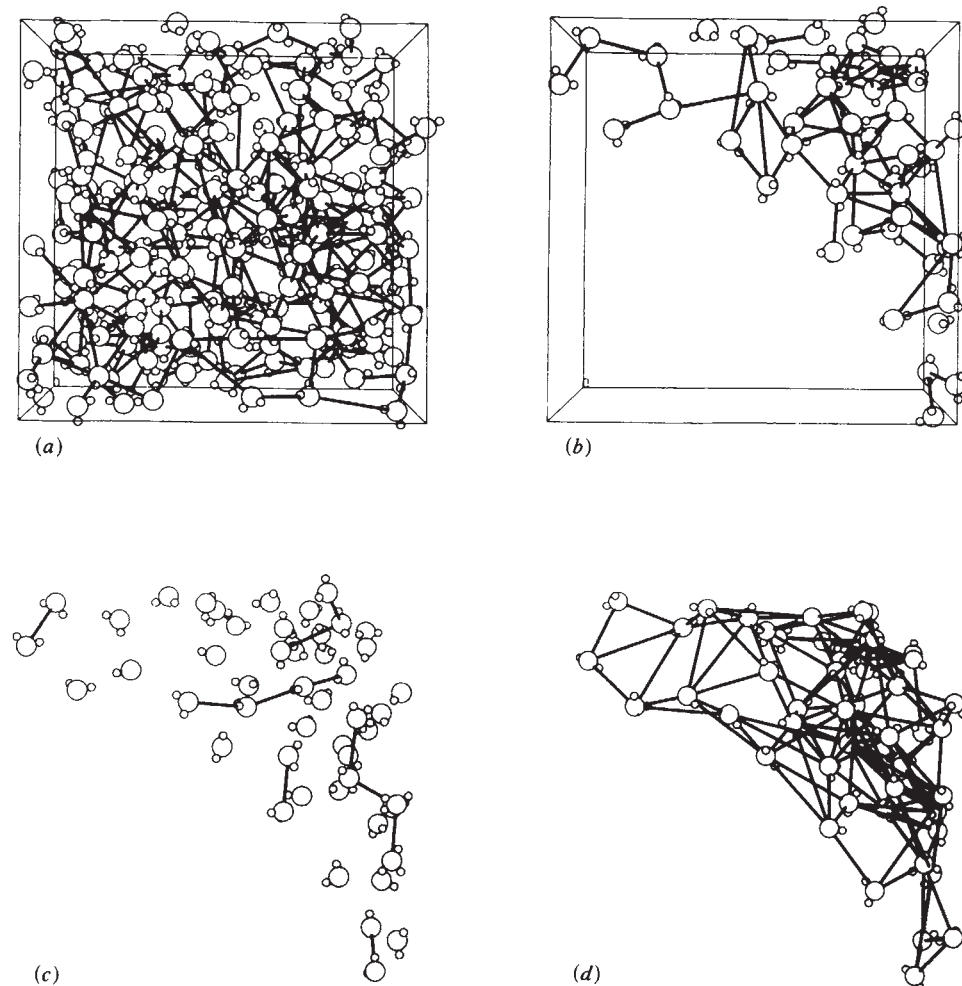


Fig. 3 Computer-drawn models of PE-simulated bulk water (a) and a cluster of 50 water molecules (b-d). The intermolecular links are notional hydrogen bonds chosen using an energy interaction criterion (cutoff at $-2.0 \text{ kcal mol}^{-1}$, using polarised electropole pair-interactions) to give an average coordination of 4.0 in the bulk. The 50 molecule cluster was chosen from a bulk configuration (b-bulk molecules not shown), then removed from the bulk environment (c), and then allowed to move freely under the Monte Carlo process (d). The initial isolation from the bulk drastically changes the bonding patterns from that of b to c, strikingly demonstrating the effects of cooperativity at an interface. Pair-additive models would be unable to distinguish between the bonding of b and c.

Ih (ref. 40), in which the hydrogen positions would be ordered. Such an ordered phase would have a predicted potential energy marginally higher ($0.4 \text{ kcal mol}^{-1}$) than the disordered phase. Although the difference is small ($\sim kT$), this alliance of increases in both entropy and binding energy is unusual. The high stability of the high pressure-hydrogen ordered ice VIII is also shown⁴¹ by the PE model to be almost wholly due to dipole-quadrupole interactions, underlining again the need for the correct sign of the quadrupole moment.

Liquid water

We have begun to examine some consequences of non-pair-additivity for the liquid phase of 'PE water'. We stress the preliminary nature of these results, in particular the short chain length so far generated (see legend, Fig. 1a), and that the model is being refined. Nevertheless, some of the results make very interesting comparisons with those from standard pair-additive potentials.

The reconstructed instantaneous configuration shown in Fig. 3a is consistent with a picture of the liquid as a non-crystalline disordered network. Closer examination, however, shows considerably lower tetrahedrality of local coordination than found in ST2, BNS or Rowlinson-water^{5,6,11,44}. Although surprising at first, this result is consistent with both the degree of tetrahedrality indicated in crystal hydrate studies⁴⁵ and the less structured picture of water given by recent neutron diffraction data^{14,15} (Fig. 1b).

Figure 1 compares radial distribution functions of PE water with experimental X-ray⁴⁶ and neutron^{14,15} results. There are evident deficiencies in the PE results (as there are deficiencies with pair-additive calculations⁸) which suggest, among other

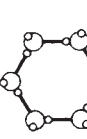
things, an inadequate treatment of the short-range repulsion; the improved agreement with the 50 °C curve suggests that our notional temperature is too low for the potential (repulsive part) used. The comparison with very recent neutron data is particularly interesting (Fig. 1b). Whereas ST2 and CF (ref. 51) both predict too much close-range structure (a consequence of over-tetrahedrality), the PE results are much more consistent with experiment. This less-structured picture is also borne out by studies²² of the distribution of effective pair interaction energies as seen in Fig. 2a. The lack of definite bimodal distribution on the negative energy side should be compared with the definite maximum observed for ST2 and BNS water⁴⁶ (Fig. 2b) which was used to support the idea of a well-defined hydrogen bond. In contrast, the PE model is consistent with a picture of a continuously varying hydrogen-bond energy rather than formation (or breaking) of discrete hydrogen bonds. This suggests that it may be dangerous to identify hydrogen bonds from the purely geometrical considerations of distances and angles. In some cases, a second-nearest neighbour geometrically is found to be a first-nearest neighbour energetically and vice versa.

Table 2 Water dimer predictions

Model	DBP*	HMS†	D‡	PKC§	CI	PE¶	Experiment ²⁷
Energy minimum (kcal mol ⁻¹)	-6.09	-4.72	-4.83	-0.460	-5.87	-5.00	
θ (deg)	122	140	140-145	135-150	143	128	$122^\circ \pm 10$
γ (deg)	51	53	52	52	48	57	$51^\circ \pm 10$
r (Å)	2.73	3.00	3.00	3.00	2.98	3.0	2.98 ± 0.04

*Refs 18, 36 predictions (quantum mechanical). †Ref. 17. ‡Ref. 28. §Ref. 43. ||Ref. 7. ¶Refs 21, 42. The angles θ , γ refer to those between the HOH bisector and the O-O separation r .

Table 3 Water cluster energies and non-pair additives (kcal mol⁻¹)

	Water cluster	Quantum mechanical	Polarisable electropole (PE)
DBP-trimers			
	Double donor	-9.33 (+1.87)	-13.58 (+0.24)
	Double acceptor	-10.14 (+2.05)	-13.66 (+0.55)
	Sequential	-14.66 (-2.06)	-16.77 (-0.83)
	Polar-cyclic	-16.83	+10.51 (-1.74)
DBP-tetramers			
	Open	-23.45	-26.02 (-1.8)
	Closed	-24.55	-27.25 (-2.2)
	Polar-cyclic	-37.91	-32.62 (-7.55)
DBP-pentamers			
	Semi-open		-36.25 (-3.12)
	Closed	-41.39	-43.70 (-5.24)
	Polar-cyclic	-53.15	-73.61 (-16.75)
DBP-hexamers			
	Polar-cyclic	-64.53	-93.67 (-23.66)

Non-pair additives are given in parentheses. The DBP-cluster configurations and the quantum-mechanical values are those given by Del Bene and Pople^{18,36}. The PE values contain just the electrostatic portion—the Lennard-Jones contribution is irrelevant to non-additivity.

The distribution of instantaneous dipole moments yields an average (enhanced) value of 2.50 D with a half-width spread of about 0.7 D. The relevance of cooperativity to dielectric, and hence solution effects, is strongly underlined. The internal energy is -6.8 kcal mol⁻¹, of which -9.0 kcal mol⁻¹ is electrostatic. The comparison with experiment³⁸ (-9.9 kcal mol⁻¹) is not good, but the above discussions of notional temperature and the sensitivity to the repulsive core should be borne in mind.

The significance of non-pair additivity at a free water interface is underlined by a Monte Carlo calculation in which a droplet of 50 water molecule centres (chosen with a large surface area) was instantaneously removed from a bulk simulated water configuration. The evolution of the removed cluster is shown in Fig. 3b-d. Note that the initial destruction of hydrogen-bonding interactions depends solely on removal of the remaining aqueous environment, and the rapid re-establishment of a new hydrogen-bonding scheme as the droplet adjusts to its isolation. Pair-additive potentials would be unable to differentiate between the bonding in the first two states (Fig. 3b, c). Hence we stress the inadequacy of using pair-additive potentials in looking at any aqueous interface, such as protein-water interactions^{47,48}.

The effects of truncating the energy calculation at a short distance, normally 7-9 Å, have been widely discussed for dipolar fluids^{49,50}. The concept of a reaction field, whereby the back-reaction of the forces from the fluid outside the cut-off radius are considered, has been proposed^{52,53} as a possible way round the problem. However, all these discussions have been within the framework of pair-additive potentials. When non-pair-additive effects are considered, there are significant changes in the nature of the problem.

Figure 4a shows the correction necessary to the energy of the 'water' assembly using different cut-offs, with and without polarisability. Clearly a premature truncation gives rise to a far

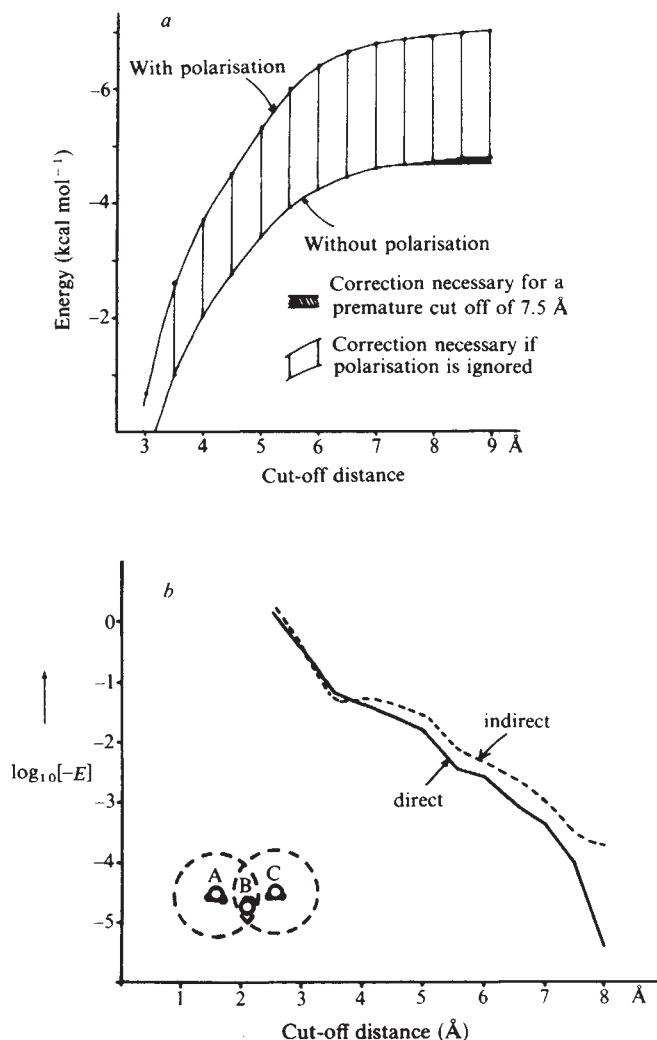


Fig. 4 a, The total energy of simulated PE-water as a function of the cut-off distance, with and without polarisation. The finely shaded portion represents the errors arising from using a cut-off distance of 7.5 Å (a common value), whereas the coarser shading indicates the errors arising from ignoring the cooperative effects of polarisation. Ignoring polarisation is clearly the dominant error. b, The variation with separation (r) of direct (—) and indirect (---) interaction energies (E) between pairs of PE-water molecules in a bulk simulation. The indirect interactions arise from mutual polarisations through intervening molecules as illustrated by the three-molecule motif; the two molecules A and C (possessing cut-off spheres indicated by the larger circles) are separated so that no direct interaction exists between them, but a third molecule B common to both spheres provides, via mutual polarisation, an indirect link between A and C. Note that the indirect (A-C type) interactions overhaul the direct (A-B, B-C type) interactions for separations above 4 Å. This probing of long-range effects cannot be realised using pair-additive potentials.

smaller error in energy than does the neglect of the non-pair-additivity. A more subtle—and perhaps more dramatic—point is that within the framework of a non-pair-additive potential, we do in fact probe extensively those interactions beyond the cut-off radius via intermediate (polarised) molecules.

Figure 4b illustrates the point with three water molecules, A, B, C, placed so that A and C lie outside each other's formal cut-off distance. However, although A and C are considered to be unable to interact directly, the polarisation effects via molecule B provide an indirect link. We now define an indirect AC interaction as the difference between the potential energy of A with and without C present: clearly, only non-additive models will give non-zero values for this indirect contribution. Figure 4b compares the direct and indirect interactions as a function of the cut-off radius, and shows in our bulk water calculation that the indirect interactions are generally of greater consequence than direct interactions, even at the larger cut-offs.

Conclusions

It is necessary to abandon the use of pair potentials for aqueous systems. Although a parametrised effective pair potential reproduces 'homogeneous' properties such as the internal energy, we would have little confidence in its ability to predict properties depending on local effects, for example, the dielectric

Received 13 June; accepted 18 September 1979.

1. Wood, W. W. in *Physics of Simple Liquids* (eds Temperley, H. N. V., Rowlinson, J. S. & Rushbrooke, G. S.) 115–230 (Wiley, New York, 1968).
2. Alder, B. J. & Wainwright, J. W. *J. chem. Phys.* **31**, 459–466 (1959).
3. McDonald, I. R. & Singer, K. Q. *Rev. chem. Soc.* **24**, 238–262 (1970).
4. Watts, R. O. & McGee, I. J. *Liquid State Chemical Physics* (Wiley-Interscience, New York, 1976).
5. Barker, J. A. & Watts, R. O. *Chem. phys. Lett.* **3**, 144–145 (1969).
6. Rahman, A. & Stillinger, F. H. *J. chem. Phys.* **55**, 3336–3359 (1971).
7. Lie, G. C., Clementi, E. & Yoshimine, M. *J. chem. Phys.* **64**, 2314–2323 (1976).
8. Swaminathan, S. & Beveridge, D. L. *J. Am. chem. Soc.* **99**, 8391–8398 (1977).
9. Lemberg, H. L. & Stillinger, F. H. *J. chem. Phys.* **62**, 1677–1690 (1975).
10. Frank, H. S. & Wen, W.-Y. *Discuss. Faraday Soc.* **24**, 133–140 (1957).
11. Stillinger, F. H. & Rahman, A. *J. chem. Phys.* **60**, 1545–1557 (1974).
12. Ben-Naim, A. & Stillinger, F. H. in *Water and Aqueous Solutions* (ed. Horne, R. A.) 295–330 (Wiley-Interscience, New York, 1972).
13. McDonald, I. R. & Klein, M. L. *Faraday Discuss. chem. Soc.* **66**, 48–57 (1978).
14. Gibson, I. P. thesis, Univ. Kent (1978).
15. Gibson, I. P. & Dore, J. C. (in preparation).
16. Arnett, E. M., Chawla, B. & Horning, N. J. *J. Sol. Chem.* **6**, 781–818 (1977).
17. Hankins, D., Moskowitz, J. W. & Stillinger, F. H. *J. chem. Phys.* **53**, 4544–4554 (1970).
18. Del Bene, J. E. & Pople, J. A. *J. chem. Phys.* **58**, 3605–3608 (1973).
19. Shepard, A. C., Beers, Y., Klein, G. P. & Rothman, L. S. *J. chem. Phys.* **59**, 2254–2259 (1973).
20. Whalley, E. *Chem. phys. Lett.* **53**, 449–451 (1978).
21. Barnes, P. in *Rep. of CECAM Workshop on Molecular Dynamics and Monte Carlo Calculations on Water* (ed. Berendsen, H. J. C.) 77–89 (CECAM, Orsay, France, 1972).
22. Barnes, P. in *Progress in Liquid Physics* (ed. Croxton, C. A.) 391–428 (Wiley, Chichester, 1978).
23. Buckingham, A. D. *Q. Rev. chem. Soc.* **13**, 183–214 (1959).
24. Neumann, D. & Moskowitz, J. W. *J. chem. Phys.* **49**, 2056–2070 (1968).
25. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water*, 16 (Clarendon, Oxford, 1969).
26. Glaeser, R. M. & Coulson, C. A. *Trans. Faraday Soc.* **61**, 389–391 (1965).
27. Dyke, T. R., Mack, K. M. & Muentner, J. S. *J. chem. Phys.* **66**, 498–510 (1977).
28. Dierksen, G. H. F. *Theoret. chim. Acta (Berl.)* **21**, 335–367 (1971).

constant which depends significantly on the enhanced dipole moment. For interfaces, the removal of even a small part of the homogeneous fluid, for example by the presence of a solute molecule or a protein surface, will upset the indirect interactions severely (compare Fig. 3). We might draw a comparison with liquid metals, where we can devise an effective pair potential for a given density; change the density, and we change the effective potential. A bulk-derived potential would not be used for a liquid metal surface or for gas. Water should be thought of in a similar way. Interface calculations, be they simple solutions or protein crystals, based on pair-additive potentials would seem very premature.

Our approach to handling cooperative effects is probably only one of several which could be useful^{54,55} and is clearly capable of further improvement. We would urge more effort be made towards developing alternative approaches, rather than continuing with applying known inadequate potentials^{13,15} to more complex systems.

We thank Dr C. Moser of CECAM, under whose auspices part of this work was carried out, Dr J. Dore for neutron data, Dr B. C. Whalley and Dr C. J. Wormald for helpful discussions. J.E.Q. is supported under SRC grant GR/A 50368. We acknowledge the computing facilities of ULCC, SRC (Daresbury), and Birkbeck College.

29. Rowlinson, J. S. *Trans. Faraday Soc.* **47**, 120–129 (1951).
30. Finney, J. L. *Faraday Discuss. chem. Soc.* **66**, 80, 86 (1978).
31. Barnes, P., Finney, J. L. & Quinn, J. (in preparation).
32. Stockmayer, W. H. *J. chem. Phys.* **9**, 398–402 (1941).
33. McCullough, J. P., Pennington, R. E. & Waddington, G. J. *Am. chem. Soc.* **74**, 4439–4442 (1952).
34. Kell, L. S., McLaurin, G. E. & Whalley, E. *J. chem. Phys.* **48**, 3805–3813 (1968).
35. McDonald, I. R. & Klein, M. L. *J. chem. Phys.* **68**, 4875–4877 (1978).
36. Del Bene, J. & Pople, J. A. *J. chem. Phys.* **52**, 4858–4866 (1970).
37. Coulson, C. A. & Eisenberg, D. *Proc. R. Soc. A* **A291**, 445–449 (1966).
38. Sharp, W. E. *UCRL Tech. Rep. No.* 7118 (1962).
39. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Clarendon, Oxford, 1969).
40. Barnes, P. in *Progress in Liquid Physics* (ed. Croxton, C. A.) 405 (Wiley, Chichester, 1978).
41. Barnes, P., Bliss, D., Finney, J. L. & Quinn, J. (in preparation).
42. Barnes, P. *Internal Report* (Crystallography Dept, Birkbeck College, 1972).
43. Popkie, H., Kistenmacher, H. & Clementi, E. *J. chem. Phys.* **59**, 1325–1336 (1973).
44. Rahman, A. & Stillinger, F. H. *J. Am. chem. Soc.* **95**, 7943–7948 (1973).
45. Olovsson, I. & Jönsson, P.-G. in *The Hydrogen Bond* (eds Schuster, P., Zundel, G. & Sandorfy, C.) Vol. 2, 393–456 (North-Holland, Amsterdam, 1976).
46. Stillinger, F. H. & Rahman, A. *J. chem. Phys.* **57**, 1281–1292 (1972).
47. Hermans, J. & Rahman, A. in *Models for Protein Dynamics* (ed. Berendsen, H. J. C.) 153–158 (CECAM, Orsay, 1978).
48. Hagler, A. T. & Moul, J. *Nature* **272**, 222–226 (1978).
49. Finney, J. L. *J. comput. Phys.* **28**, 92–102 (1978).
50. Barker, J. A. & Watts, R. O. *Molec. Phys.* **26**, 789–792 (1973).
51. Stillinger, F. H. & Rahman, A. *J. chem. Phys.* **68**, 666–670 (1978).
52. Watts, R. O. *Molec. Phys.* **28**, 1069–1083 (1974).
53. Van Gunsteren, W. F., Berendsen, H. J. C. & Rullmann, J. A. C. *Faraday Discuss. chem. Soc.* (in the press).
54. Berendsen, H. J. C. & van der Velde, G. A. in *Rep. CECAM Workshop on Molecular Dynamics and Monte Carlo Calculations on Water* (ed. Berendsen, H. J. C.) 63–71 (CECAM, Orsay, 1972).
55. Stillinger, F. H. & David, C. S. *J. chem. Phys.* **69**, 1473–1484 (1978).
56. Le Ferre, E. J., Nightingale, M. R. & Rose, J. W. *J. mech. Engng Sci.* **17**, 243–251 (1975).
57. Lie, G. C. & Clementi, E. *J. chem. Phys.* **64**, 5308–5309 (1976).

Subduction, the geoid, and lower mantle convection

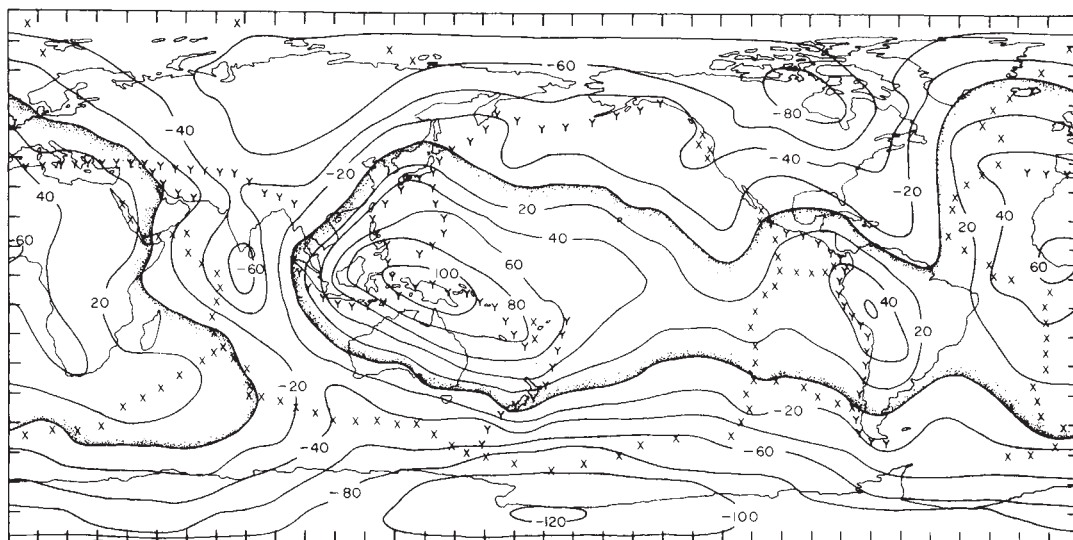
Clement G. Chase

Department of Geology and Geophysics, University of Minnesota, Minneapolis, Minnesota 55455

Geoid models show that net uncompensated masses in subducting lithosphere are less than thermal models predict, and do not require elevation of mantle phase changes in the slabs. The largest deviations from hydrostatic equilibrium ignore plate boundaries and may be ascribed to lower mantle convection uncoupled to plate motions.

THE EARTH'S gravity field gives the most direct evidence on the distribution of mass in its interior. In particular, areas of anomalous density associated with thermal convection—cold subducting slabs, mantle upwellings, deflection of internal density boundaries—should be reflected in the global gravity potential. Some surface tectonic features have characteristic large scale gravity anomalies^{1,2}, especially the oceanic subduction zones with their associated geoid (Fig. 1) and gravity

Fig. 1 Contours of the GEM8 geoid relative to a hydrostatic figure of flattening 1/299.75. In all the figures: solid contours are at 20-m intervals; dashed contours at intermediate 10-m intervals; uphill side of the zero contour stippled. Geoids were synthesised from coefficients of degree 2–20. Y, point on subduction zone; X, point on spreading centre. Cylindrical equidistant projection.



highs. However, most of the observed variation in the gravity field is not closely related to plate boundaries. Here forward modelling is used to match the observed field, thereby placing limits on the size and depth of some of the anomalous regions. Most previous quantitative interpretations of global gravity^{3–9} have used direct or statistical inversion of the gravity potential.

The geoid

It proved advantageous to work with the geoid rather than with global gravity anomalies, which amplify the higher degrees of the spherical harmonic expansion of the potential. In geopotential models those coefficients are not as well determined as the lower degrees. Thus the geoid, which weights the harmonics evenly, is less subject to errors of measurement. The geoid also 'sees' deeper than the gravity field, because the effect of a mass on the geoid decreases directly with depth, while the gravity effect decreases with the square of the depth. Because the geoid is not as sensitive to the exact depth of perturbing bodies, their masses can be estimated more reliably without knowing the details of their vertical distribution.

All departures from hydrostatic equilibrium within the Earth must be supported by the strength of rocks or by viscous stresses, so the appropriate geoidal reference surface for tectonic purposes is the hydrostatic rotational figure of the Earth. Figure 1 shows the elevation of the nonhydrostatic geoid for GEM8¹⁰. The shortest wavelength represented is ~1,000 km.

Subduction zone model

It seems to be accepted that excess (uncompensated) mass of the subducting slabs is important in causing oceanic slabs to move (see ref. 11 and refs therein). Watts and Talwani¹² inferred from gravity models of outer rises that the cold slab was somehow regionally compensated. However, Grow and Bowin¹³ have criticised their method and conclusion. Uncompensated slabs are used here since large averaging depth causes the geoid to respond most strongly to net excess or deficit of mass.

The starting point was the thermal slab model of McKenzie¹⁴ as modified for finite slab length by Molnar and Gray¹⁵. The slab's down-dip length was taken as length of the seismic zone¹⁶, except for Patagonia and the continental collision zones, where slabs to 200 km were assumed. Approximating points (Figs 1, 2) were then assigned masses determined by the local subduction rate from a global relative motion model¹⁷ and the area of slab they represent. Behind-arc spreading was not taken into account. Depths of the point masses are shown in Table 1. Spherical harmonic coefficients of the point mass potentials¹⁸ were then calculated and summed.

The simple thermal models markedly overestimate the net excess mass localised in most subducting slabs, turning positive anomalies into negative ones. An empirical subduction zone model (Fig. 2) was derived by independently adjusting the amount of excess mass for each subduction zone, usually downwards (Table 1), until the residual geoid (Fig. 3) showed as

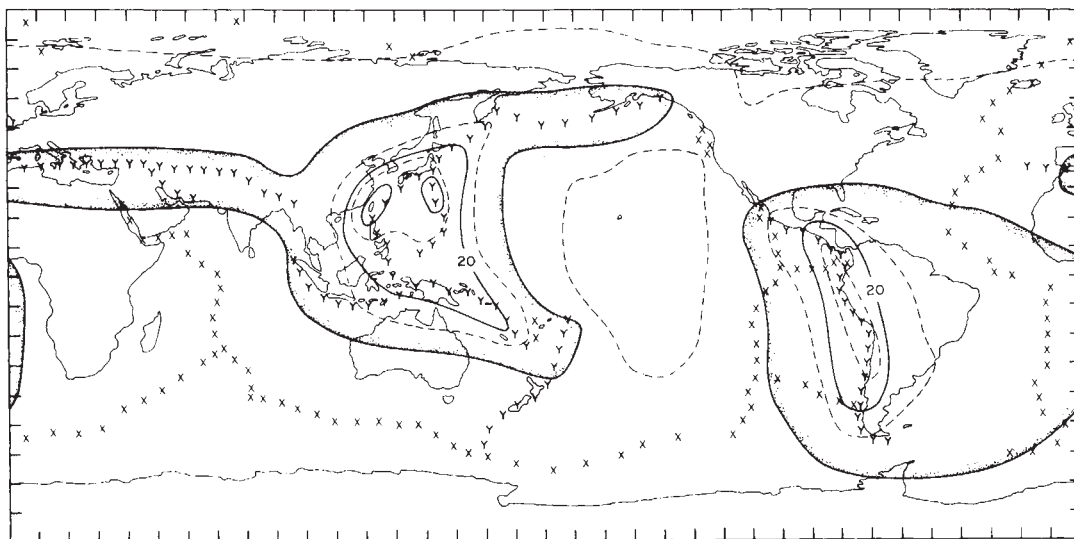


Fig. 2 Model of geoidal undulations due to excess uncompensated masses localised in cold subducting slabs (Table 1). Note the broad, low amplitude high in the South Atlantic. This is a 'farside' effect caused by displacement of the centre of mass towards the subduction zones in the Pacific. Centre of mass offset is 16 m towards 23° N, 165° E.

Table 1 Subduction zone (cold slab) and deep source models

Feature	Depth (km)	Length (km)	Excess mass (10^{15} kg km $^{-1}$)	Cold slab theory (10^{15} kg km $^{-1}$)*
Subduction zones				
Middle America	200	2,864	1.10	1.03
Peru–Chile	325	6,383	1.08	3.60
Patagonia	150	1,601	0.26	0.18
Tonga–Kermadec	350	3,708	0.41	2.37
New Hebrides–New Guinea	325	5,129	1.24	1.38
Java	350	5,384	0.37	2.35
Mariana–Bonin	375	3,498	1.33	2.29
Philippine	325	2,989	0.77	1.49
Ryukyu	325	2,215	1.62	0.89
Honshu	325	966	1.11	4.25
Kuriles	300	2,083	0.64	2.64
Aleutians	250	4,002	0.70	1.40
Mediterranean	250	5,608	0.24	0.20
Zagros	150	2,256	0.22	0.45
Himalaya	150	5,196	0.27	0.36
Deep sources				
Pacific	2,000–2,900	16,863	15.8	
Africa	2,000–2,900	18,651	11.9	
India–Antarctica	1,000–1,200	17,300	–4.18	
North America	1,000	6,111	–4.30	
Siberia	1,000	1,671	–4.29	
Arabia	1,200	(only 1 point)	8.36×10^{18} kg	

* Calculated from slab model of Molnar and Gray¹⁵ using down-dip length of seismic zone from Isacks and Molnar¹⁶, lithospheric thickness 100 km, $T_1 = 1,200$ K, $\rho = 3,300$ kg m $^{-3}$, $c_p = 1,050$ J kg $^{-1}$ K $^{-1}$, $\alpha = 3 \times 10^{-5}$ K $^{-1}$, $\kappa = 3.77$ W m $^{-1}$ K $^{-1}$, convergence rates from ref. 17.

little correlation with the position of the subduction zones as possible.

The subduction zone model (Fig. 2) has positive masses only but negative anomalies in some places, and positive values in places where there are no anomalous masses. The omission of zero and first degree terms from the gravity field means that a positive excess mass will have a positive geoid anomaly directly over it, negative values at greater distance, and a gentle positive on the opposite side of the Earth. This 'farside' effect is caused by displacement of the centre of mass towards the perturbing body.

There may be several reasons that the final model required less mass than theory suggested. The thermal model may not account for all important processes of heat transfer and generation. Because the topography of the oceanic trenches is probably not locally compensated, a trench can contribute up to 0.5×10^{15} kg km $^{-1}$ of mass deficit to counteract the slab's contribution. There may also be local phase boundary deflections caused by convection, and thermal anomalies beneath back-arc basins that partially compensate the cold slab. Elevation of the olivine-spinel phase transition within the cold slab^{19,20}, increasing its

excess mass, would worsen misfit of the thermal model. Sung and Burns²¹ suggest that sluggish kinetics will delay transitions in the cold heart of the slab to greater depths, reducing or reversing the effect of the phase change.

A modest amount of uncompensated slab was permissible for continental collision zones of the Himalaya, Zagros and Mediterranean areas. This is surprising considering the importance of the Himalaya in development of isostatic theory, but there is some support from surface gravity studies²².

The subduction zone model does not remove all geoidal highs located near the trenches, leaving broad positive anomalies on the continental side of the East Asian and Peru–Chile trenches (Fig. 3). These may reflect the pressure gradient necessary to drive asthenospheric counterflow from trenches to ridges. Global models of the flow^{23–25} suggest that geoidal elevation should be roughly symmetrical about most convergence zones, but the residual highs are not. However, cooling of the oceanic lithosphere can reduce the elevation on the ocean side of the trench²⁶.

Another possibility has been pointed out by F. M. Richter (personal communication). Mantle cooled by the subducting

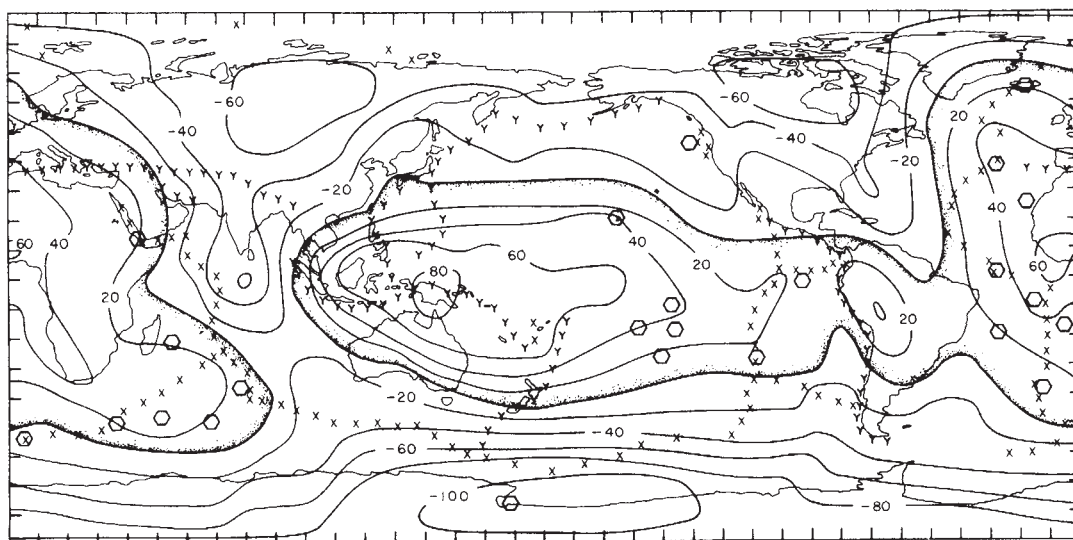
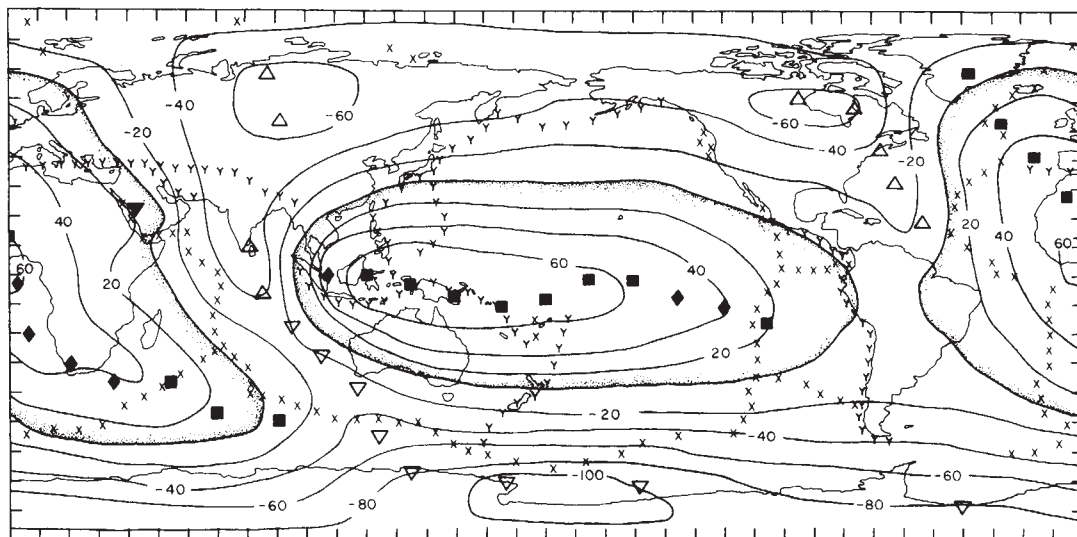


Fig. 3 Residual after subtraction of subduction zone model (Fig. 2) from observed geoid (Fig. 1). Large amplitude and wavelength anomalies remain. $\odot$, The location of selected hotspots, mostly oceanic. They cluster strongly about the positive anomalies, mostly near the +20 to +40 m contour.

Fig. 4 Model of deep uncompensated point masses designed to approximate residual of Fig. 3. Depths: Δ , 1,000 km; ∇ , 1,200 km; $\diamond$, $\blacklozenge$, 2,000 km; $\blacksquare$, $\square$, 2,900 km. Open symbols, negative point masses; filled symbols, positive point masses. Note the continuity of the anomalies and that the negative masses are all shallow. Centre of mass offset 64 m towards 6° N, 64° E.



slab cannot be removed from near the slab faster than conduction (very slow) or convection permits. This could leave a broad, fairly shallow dense area behind the subduction zone. Note that the residual highs are located where the subduction zones used to be, because all of these trenches move oceanwards in the hotspot frame of reference²⁷.

Large residual anomalies

The most remarkable features in Fig. 3 are three enormous 60–80-m amplitude geoidal anomalies. There are two positive features—one in the Pacific, stretching from Sumatra to the Galapagos, and the other from south-east of Africa to Greenland—and one discontinuous negative anomaly from Antarctica through India and Siberia to the North Atlantic. Clearly most variation of the geoid from hydrostatic equilibrium is not due to subduction zones.

These three large geoidal anomalies share several characteristics that need explanations. (1) Great size: each of the anomalies is at least 18,000 km in length, and the positive anomalies are more than 6,000 km wide, the negative is about 4,500 km wide. (2) Large amplitudes: all have amplitudes greater than 40 m for much of their length. (3) Lack of correlation to surface tectonic features: all three anomalies cross several plate boundaries with little apparent disturbance. There is no correspondence to type or age of lithosphere. (4) Strong correlation to location of oceanic hotspots: of the 24 selected hotspots marked on Fig. 3, all but three lie within the positive anomalies.

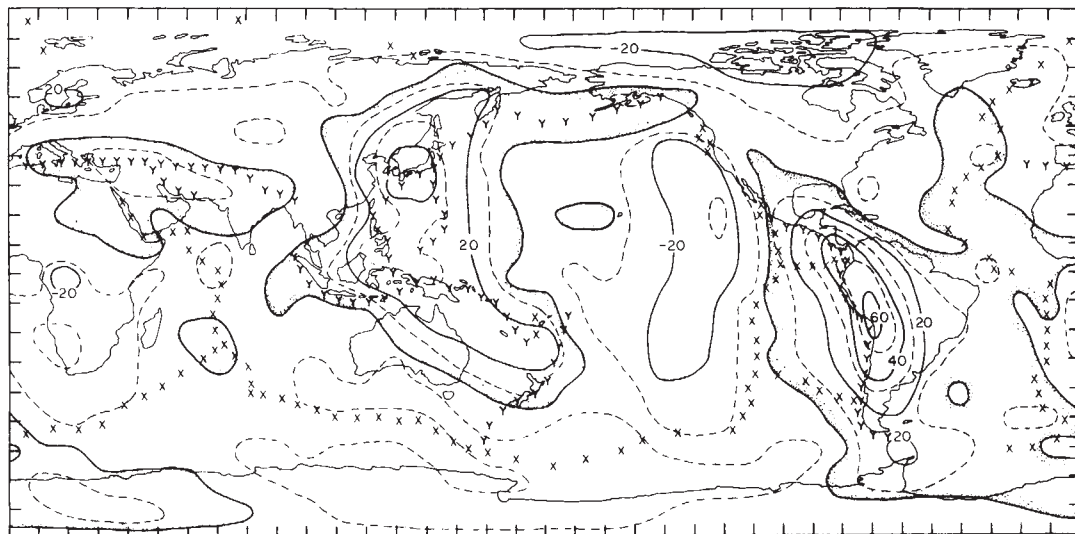
Deep source model

To find maximum source depth, I have constructed a point mass model for the large residual anomalies (Fig. 4). The point sources for the positive anomalies could be placed quite deep, at 2,000 or 2,900 km (core-mantle boundary). This depth is a maximum, because a broad, smooth anomaly can also come from a more gradual density fluctuation at shallower depth. The sources of the negative anomaly could be no deeper than 1,200 km, however. The estimate of excess mass for the features (Table 1) is a close upper limit and not very sensitive to depth of the source mass distribution. The positive anomalies, as well as being broader and possibly deeper in origin, require much more excess mass than the negative one.

This model for the non-hydrostatic part of the Earth's shape includes the excess equatorial bulge. The largest anomaly, the equatorial Pacific high, determines the maximum non-hydrostatic moment of inertia of the Earth, and therefore the position of the rotational axis²⁸. Should the African high grow larger than the Pacific high, the whole Earth would shift relative to the axis—true polar wander, with Greenland moving to the Equator!

The observed geoid with the deep source model subtracted is shown in Fig. 5. Subduction-related features are most prominent (compare Fig. 2). There are other anomalies of note: positive anomalies of 2–8 m height over the mid-ocean ridges (see ref. 29); and a 10-m high in mid-Pacific west of the Hawaiian Islands³⁰. The broad low in the eastern Pacific may be related to asthenospheric counterflow, as discussed above.

Fig. 5 Observed geoid with deep mass model of Fig. 4 removed. Compare with Fig. 2.



Lower mantle convection?

What, then, is the origin of the large anomalies (Fig. 3) that dominate the non-hydrostatic part of the Earth's geoid? Rocks are not strong enough to support these features, so convective processes must be responsible^{3,5,31}. The lack of correlation of the large anomalies to plate boundaries and motions allows a more specific interpretation of where the convection is located: it must be below the flow regime associated with plate tectonics.

Whole mantle convection closely coupled to surface plate motions has been invoked lately^{25,32-34}, but is clearly not permissible if the large geoidal anomalies unambiguously reflect the location of lower mantle convection cells. Such a geometry could not satisfy mass balance requirements for plate motions²⁴ and there is no particular relationship to plate motions or interplate stresses¹¹. For example, convection associated with the Pacific high should either extend or compress the Pacific plate in a north-south direction, neither of which is observed¹¹. Thus the residual geoidal anomalies are direct evidence favouring the model advanced by McKenzie and Weiss³⁵ of two layer convection with upper and lower mantle circulating separately.

It is not obvious whether the positive anomalies would mark upwellings or downwellings in the lower mantle convection. Analytical³⁶ and numerical³⁷ models of mantle convection suggest that deflections of density boundaries will dominate the gravity field in many cases, leading to positive anomalies over upwellings. The tendency of hotspots to occur over geoidal highs may indicate hotter than normal material below. The hotspots themselves imply some interchange of material between lower and upper mantle, if they do originate within a lower convective system. Global heat flow patterns³⁸ show little correlation with the large geoidal anomalies. Seismic travel times under the North Atlantic³⁹ resemble the geoidal patterns there, slow travel times corresponding to geoidal lows. Globally, deep travel time anomalies⁴⁰ also show some correlation with geoid anomalies in the low degree harmonics. A great deal more work needs to be done on the intriguing possibility that the Earth's gravity field reveals convective patterns hidden at depth.

This work was supported by NSF grant EAR 75-21622. I thank the Earth Science Board, University of California, Santa Cruz for hospitality during my sabbatical when this work was done, and B. Hager and G. Davies for helpful comments.

Received 11 June; accepted 27 September 1979.

1. Kaula, W. M. in *The Nature of the Solid Earth* (ed. Robertson, E. C.) 385-405 (McGraw-Hill, New York, 1972).
2. Kaula, W. M. *Tectonophysics* **13**, 341-359 (1972).
3. Runcorn, S. K. *Geophys. J. R. astr. Soc.* **14**, 375-384 (1967).
4. Hide, R., & Horai, K. *Phys. Earth planet. Int.* **1**, 305-308 (1968).
5. Runcorn, S. K. *Tectonophysics* **13**, 623-637 (1972).
6. Higbie, J. W. & Stacey, F. D. *Nature phys. Sci.* **234**, 130-132 (1971).
7. Bott, M. H. P. *Earth planet. Sci. Lett.* **11**, 28-34 (1971).
8. McQueen, H. W. S. & Stacey, F. D. *Tectonophysics* **34**, T1-T8 (1976).
9. Khan, M. A. *Geophys. J. R. astr. Soc.* **48**, 197-209 (1977).
10. Wagner, C. A., Lerch, F. J., Brown, J. E. & Richardson, J. A. *J. geophys. Res.* **82**, 901-914 (1977).
11. Richardson, R. M., Solomon, S. C. & Sleep, N. H. *Rev. Geophys. Space Phys.* **17**, 981-1019 (1979).
12. Watts, A. B. & Talwani, M. *Bull. geol. Soc. Am.* **86**, 1-4 (1975).
13. Grow, J. A. & Bowin, C. O. *J. geophys. Res.* **80**, 1449-1458 (1975).
14. McKenzie, D. P. *Geophys. J. R. astr. Soc.* **18**, 1-32 (1969).
15. Molnar, P. & Gray, D. *Geology* **7**, 58-62 (1979).
16. Isacks, B. & Molnar, P. *Rev. geophys. Space Phys.* **9**, 103-174 (1971).
17. Chase, C. G. *Earth planet. Sci. Lett.* **37**, 355-368 (1978).
18. Pollack, H. N. *J. geophys. Res.* **78**, 1760-1768 (1973).

19. Smith, A. T. & Toksoz, M. N. *Geophys. J. R. astr. Soc.* **29**, 289-318 (1972).
20. Schubert, G., Yuen, D. A. & Turcotte, D. L. *Geophys. J. R. astr. Soc.* **42**, 705-735 (1975).
21. Sung, C. M. & Burns, R. G. *Tectonophysics* **31**, 1-32 (1976).
22. Kono, M. *Geophys. J. R. astr. Soc.* **39**, 283-299 (1974).
23. Harper, J. F. *Geophys. J. R. astr. Soc.* **55**, 87-110 (1978).
24. Chase, C. G. *Geophys. J. R. astr. Soc.* **56**, 1-18 (1979).
25. Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **84**, 1031-1048 (1979).
26. Schubert, G., Yuen, D. A., Froidevaux, C., Fleitout, L. & Souriau, M. *J. geophys. Res.* **83**, 745-758 (1978).
27. Chase, C. G. *J. geophys. Res.* **83**, 5385-5387 (1978).
28. Goldreich, P. & Toomre, A. *J. geophys. Res.* **74**, 2555-2567 (1969).
29. Haxby, W. F. & Turcotte, D. L. *J. geophys. Res.* **83**, 5473-5478 (1978).
30. Watts, A. B. & Leeds, A. R. *Geophys. J. R. astr. Soc.* **50**, 249-277 (1977).
31. Gough, D. I. *Geol. Soc. Can. Prog. w. Abstr.* **2**, 21 (1977).
32. O'Connell, R. J. *Tectonophysics* **38**, 119-136 (1976).
33. Davies, G. F. *Geophys. J. R. astr. Soc.* **49**, 459-486 (1977).
34. Elsasser, W. M., Olson, P. & Marsh, B. D. *J. geophys. Res.* **84**, 147-155 (1979).
35. McKenzie, D. & Weiss, N. *Geophys. J. R. astr. Soc.* **42**, 131-174 (1975).
36. Pekeris, C. L. *Mon. Not. R. astr. Soc. geophys. Suppl.* **3**, 343-367 (1935).
37. McKenzie, D. P., Roberts, J. M. & Weiss, N. O. *J. Fluid Mech.* **62**, 465-538 (1974).
38. Chapman, D. S. & Pollack, H. N. *Earth planet. Sci. Lett.* **28**, 23-32 (1975).
39. Stewart, I. C. F. *Geophys. J. R. astr. Soc.* **49**, 487-497 (1977).
40. Dziewonski, A., Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **82**, 239-255 (1977).

Specificity of the antibody response of rabbits to a self-antigen

Ronald Jemmerson* & Emanuel Margoliash

Department of Biochemistry and Molecular Biology, Northwestern University, Evanston, Illinois 60201

The demonstration that the antigenic determinants on rabbit cytochrome c that elicit antibodies in rabbits occur in regions of variability among mammalian cytochromes c suggests the existence of an evolutionary process which eliminates genetic specificities for self-determinants.

DESPITE the considerable attention which has been given to the antigenic properties of proteins, we still do not know why certain parts of a protein antigen elicit antibody responses while other regions remain immunologically silent. As the B-cell specificity repertoire that is expressed is limited¹, it follows that not every molecular substructure of an antigen would have a corresponding B cell reactive to it. But what restraints act to limit the B-cell specificity repertoire?

One theory is that, because of the phenomenon of immunological tolerance, antibodies may not be elicited against

those regions of an antigen which share structural or chemical properties with self-molecules of the host. Indeed, studies on the antigenicity of cytochrome c indicate that the regions of amino acid sequence difference between the cytochrome c immunogen and the corresponding host's protein are antigenic²⁻⁵. For example, one of the six residue differences between horse and rabbit cytochromes c is residue 60, where horse cytochrome c has lysine and the rabbit protein has glycine. These two cytochromes c have 18 other lysyl residues in common. When rabbits are immunised against horse cytochrome c, the antibodies directed against the region of lysine 60 comprise 20-40% of the total antibody and the other lysyl residues do not elicit antibody responses^{4,5}. It is difficult to explain this observation other than by tolerance. The concept of a particular chemistry inducing antigenicity must be excluded in this instance as only one of 19 lysine residues is antigenic.

Although immunological self-tolerance seems to be a dominant factor affecting the specificity of the immune response to cytochrome c, it was early observed that not only did antibodies elicited against a heterologous cytochrome c bind the rabbit protein, but also when polymerised with glutaraldehyde⁶

* Present address: Department of Cellular and Developmental Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92037.

or covalently coupled to bovine γ -globulin⁷, rabbit cytochrome *c* itself was immunogenic in rabbits. The finding that antibodies elicited against a foreign cytochrome *c* cross-reacted with rabbit cytochrome *c* is attributed to the more stringent requirements for stimulating a B cell than for binding of an antigen to the cell receptor (antibody)^{8,9}. Thus, the observation is not inconsistent with the existence of a state of tolerance in rabbits to rabbit cytochrome *c*. However, the finding that rabbit cytochrome *c*, when in certain molecular forms, does initiate an antibody response is not as readily understood. The quantitatively lower response evoked by the self-antigen, as opposed to heterologous cytochromes *c*, indicates a degree of immunological tolerance to rabbit cytochrome *c* and it may be that modification by coupling to a carrier or by polymerisation may have enabled the protein to trigger a response that would normally not be allowed.

In any case, the ability to elicit an antibody response to a self-antigen indicates that immunological self-recognition is not the only mechanism allowing certain sites on an antigen to remain immunologically silent. We considered that by establishing which sites on rabbit cytochrome *c* elicit antibody responses in rabbits, we might elucidate factors other than lack of self-recognition that would influence the antigenicity of only certain regions of a molecule.

To obtain autoantibodies against rabbit cytochrome *c*, eight adult New Zealand white male rabbits were injected three times over 6 weeks with the glutaraldehyde-polymerised form of the protein (~9 mg total) emulsified in Freund's complete or incomplete adjuvants. The cytochromes *c* used in this study and the glutaraldehyde-polymerised rabbit cytochrome *c* were prepared as previously described³. Rabbit cytochrome *c* was chemically fragmented using cyanogen bromide and the peptides (residues 1–65, 66–80 and 81–104) were separated according to the procedure of Corradin and Harbury¹⁰. Other methods used, including antibody fluorescence quenching titration and modified Farr assay, have been reported elsewhere³.

Characterisation of the antibody response

All eight rabbits immunised with rabbit cytochrome *c* polymers produced antibodies specific for the monomer, with serum titres between 115 and 230 μ g antibody ml⁻¹ as determined by immunoadsorption of the antibodies on rabbit cytochrome *c* coupled to CNBr-activated Sepharose 4B. The serum titres were 5–20% of the response observed against heterologous cytochromes *c* differing from the rabbit protein at 6–9 residues out of 104. Clearly, rabbits display a degree of 'tolerance' to their own cytochrome *c*.

The rabbit anti-rabbit cytochrome *c* antibodies adsorbed on the insolubilised antigen had an apparent molecular weight of 150,000 based on gel filtration studies on a calibrated column of Sephadex G-150, migrated electrophoretically as a single band

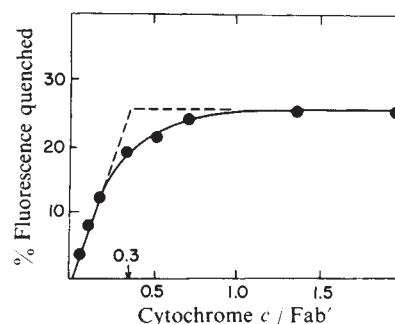


Fig. 1 Fluorescence quenching titration of rabbit anti-rabbit cytochrome *c* Fab' with rabbit cytochrome *c*. The antibodies were isolated from pooled antisera on insolubilised rabbit cytochrome *c*, digested with 2% pepsin for 8 h, then reduced with cysteine and alkylated with iodoacetamide.

in the γ region, and showed precipitation in agarose film with goat anti-rabbit γ -chain-specific but not μ -chain-specific antisera. Thus, the antibodies elicited in rabbits against their own protein were IgG. The stoichiometry of the antibody-binding reaction, as determined by fluorescence quenching assays between the Fab' fragments prepared from the antibodies and rabbit cytochrome *c*, was 3:1, indicating the presence of at least three separate antigenic determinants on rabbit cytochrome *c* (Fig. 1).

Modified Farr assays of rabbit anti-rabbit cytochrome *c* sera showed that the antibodies distinguished between rabbit and heterologous cytochromes *c* that differed at two to six amino acid residue positions (Fig. 2, Table 1). Thus, mouse cytochrome *c*, which varies from the rabbit protein at only two positions—residues 44 and 89—did not compete with ¹²⁵I-labelled rabbit cytochrome *c* for the antibody-binding sites as effectively as did rabbit cytochrome *c*. The dog protein displaced only 30% of the labelled rabbit cytochrome *c*, even though it differs from the rabbit protein at only five residue positions—44, 62, 88, 89 and 103. These results show that the rabbit anti-rabbit cytochrome *c* antibodies are directed against regions of the molecule where the rabbit and the other mammalian cytochromes *c* differ in amino acid sequence.

Identification of the sites of antibody binding

At least one antibody population bound in the carboxyl-terminal segment comprising residues 88–104. Indeed, dog cytochrome *c* was a less effective inhibitor in the modified Farr assay than was the beef protein, which differs from it only at residues 88, 92 and 103. Furthermore, the peptide consisting of residues 81–104 of rabbit cytochrome *c* inhibited the binding of approximately 20% of the labelled rabbit cytochrome *c* when present in a 20-fold molar excess over the latter (Fig. 3). The difference in binding between mouse and rabbit cytochromes *c* may indicate that the antigenic site is in the region of residue 89 as these two proteins differ only at that position and at residue 44. In any case, the data indicate that residue 44 or 89 or both are within antigenic determinants. In contrast to the 81–104 peptide, the 66–80 peptide failed to inhibit any binding of the antibodies to rabbit cytochrome *c*. However, the haem-containing 1–65 peptide displaced 80% of the labelled protein. Thus, the 1–65 and the 81–104 peptides contained the whole of the autoantigenic activity of rabbit cytochrome *c*. Further one of the antibodies binding to the 1–65 fragment was found to be directed against the region of residue 62. Thus, mouse and kangaroo cytochromes *c*, which, like the rabbit protein, carry aspartic acid at that position, were more effective inhibitors of rabbit cytochrome *c* than was guanaco cytochrome *c*, which has glutamic acid (Fig. 2). The only other position at which guanaco cytochrome *c* differs from the rabbit protein is 89, where the former

Table 1 Comparison of the amino acid sequences of mammalian cytochromes *c*

	3	3	4	4	5	6	6	8	8	9	10
Cytochrome <i>c</i>	3	5	4	7	8	0	2	8	9	2	3
Rabbit ¹²	H	L	V	S	T	G	D	K	D	A	N
Mouse ^{13,14}	—	—	A	—	—	—	—	—	G	—	—
Kangaroo ¹⁵	N	I	P	—	I	—	—	—	G	—	—
Guanaco ¹⁶	—	—	—	—	—	—	E	—	G	—	—
Beef ¹⁷	—	—	P	—	—	—	E	—	G	E	—
Horse ¹⁸	—	—	P	T	—	K	E	—	T	E	—
Dog ¹⁹	—	—	P	—	—	—	E	T	G	—	K

The single letter code for amino acids is used. Residue positions are given vertically, for example, ₃ referring to position 33. Only the residues that differ among the cytochromes *c* listed are shown. — Indicates that the residue is identical to that of the rabbit protein.

has a glycine and the latter aspartic acid. Both mouse and kangaroo cytochromes *c* are identical to the guanaco protein at position 89, so that the greater affinity of these two cytochromes *c* for the antibodies must be due to their greater similarity to the rabbit protein at position 62.

As the stoichiometry of the reaction between rabbit cytochrome *c* and its antibodies is three, in addition to the antibodies binding in the carboxyl-terminal segment and in the region of residue 62 there is at least one other antibody population. It probably binds to the amino-terminal segment of the protein represented by the 1-65 peptide, because as already noted, that peptide accounts for a large proportion of the autoantigenic activity of the molecule, and the 81-104 fragment contains another antigenic determinant accounting for the remainder of the activity.

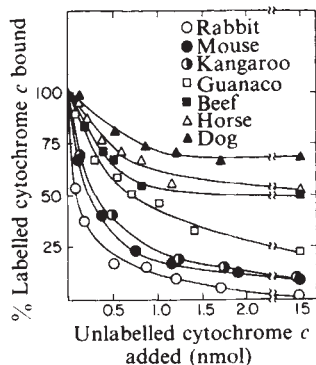


Fig. 2 Competitive binding assays between rabbit and various mammalian cytochromes *c* and antibodies elicited in rabbits against the glutaraldehyde-polymerised form of rabbit cytochrome *c*. Rabbit cytochrome *c* was iodinated with carrier-free ^{125}I by the method of Hunter and Greenwood²⁰. The antiserum used was pooled from four rabbits and diluted 1:4 with 6.2 mM sodium borate-boric acid buffered at pH 8.6, containing 130 mM sodium chloride. Each reaction mixture contained 1.36 nmol ^{125}I -labelled rabbit cytochrome *c*, 0.20 ml diluted antiserum and amounts of unlabelled cytochrome *c* ranging from 0 to 1.5 nmol. Antibody-cytochrome *c* complexes were precipitated in 50% saturated ammonium sulphate and washed three times. The amount of radioactivity precipitated was corrected for nonspecific binding using normal rabbit serum. In the absence of inhibitor 0.038 nmol of ^{125}I -labelled rabbit cytochrome *c* were precipitated.

Do specificities for self-determinants tend to be eliminated during evolution?

From the above analysis, two of the three antigenic sites on the self-antigen rabbit cytochrome *c* have been localised to the regions around residue 62 and the carboxyl-terminal segment probably around residue 89. Those two sites correspond to regions on heterologous cytochromes *c* that are antigenic in rabbits. Indeed, the antigenic sites on guanaco and mouse cytochromes *c* were found to be centred around residues 62 and 89 and residues 44, 62 and 89, respectively³, and those on horse cytochrome *c* were in the regions of residues 44, 60 and 89 or 92 (refs 4, 5). The area around residue 44 in rabbit cytochrome *c* may be the third autoantigenic site of that protein.

What property of these regions of cytochrome *c* makes them antigenic? It is unlikely that the particular amino acid side chains share an antigenic property, as each cytochrome *c* examined with respect to antigenic specificity has different amino acid residues in the immunodominant position and the antibodies elicited by each cytochrome *c* antigen are specific for its own side chains²⁻⁵. Possibly, the tertiary structure of the molecule in the antigenic regions is such that interactions with cells of the immune system occur more readily. This explanation also seems unlikely as the cytochrome *c* molecule is a tightly folded oblate spheroid and no externally located backbone sequence is par-

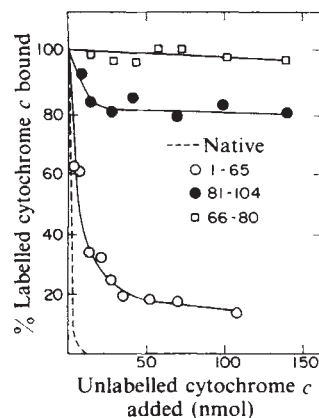


Fig. 3 Competitive binding assays between rabbit cytochrome *c* and cyanogen bromide-cleaved peptides of the protein and antibodies elicited in rabbits against rabbit cytochrome *c*. The assays were carried out as described in Fig. 2 legend.

ticularly better situated than any other for increasing the probability of extramolecular interactions. In this regard, there are no sharp bends in cytochrome *c* as there are in myoglobin that could explain the antigenicity of certain sites on that molecule¹¹.

Notwithstanding any unknown chemical features which of themselves might allow certain regions to be antigenic, it is particularly interesting that the antigenic sites on rabbit cytochrome *c* correspond to the areas on the molecule where other mammalian cytochromes *c* differ from it (Fig. 4) and in fact, a region in which all cytochromes *c* are identical (residues 66-80) was demonstrated not to be antigenic in rabbit cytochrome *c*.

A possible explanation for this is that only certain regions of a self-antigen elicit immune responsiveness because other specificities are absent from the repertoire (not carried in the germ line). Selective pressures against the maintenance of antibody specificities to self-antigens could act during evolution to eliminate specificities for self-antigenic determinants from the antibody repertoire. Because self-antigens would continue to vary structurally in the course of evolution, as selective pressures begin to effect the elimination of the corresponding antibody specificities, those specificities corresponding to the most variable regions of a self-antigen, namely, those most likely to have varied recently, are the least likely to have been eliminated. Selective pressures would be most effective against those antibodies specific for the evolutionarily conserved regions of a self-molecule because of the long exposure time required for evolutionary elimination.

Thus, if this interpretation is correct, antibodies elicited against a self-antigen should correspond to those regions of the molecule which have varied evolutionarily. That seems to be the

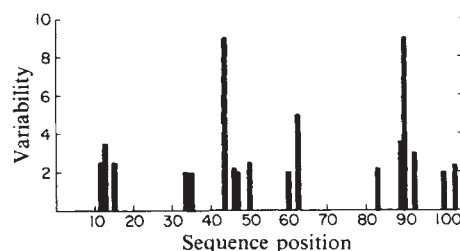


Fig. 4 Variability of mammalian cytochromes *c*. Variability is defined as the number of different amino acids at a given position divided by the frequency of the most common amino acid at that position²¹. Eighteen mammalian cytochrome *c* sequences were used for this^{22,23}. The variabilities at residues 11, 12 and 15 are attributable to amino acid substitutions in the later evolving primate cytochromes *c* and thus, the region around residues 11, 12 and 15 would not be antigenic in rabbit cytochrome *c* by the mechanism discussed in the text.

case with the antibody response in rabbits against rabbit cytochrome *c*. One could, therefore, consider that there are at least two biological levels at which an individual is rendered tolerant to self-antigens—at the developmental level, clones reactive with self are aborted or tolerated; at the evolutionary level, genes corresponding to specificities against self tend to be eliminated. Although this is not the only possible interpretation,

it would be consistent only with a germ-line or programmed somatic mutation theory of antibody diversity and cannot be reconciled with a theory of diversity involving somatic mutation alone.

We thank Drs Susan Pierce and Norman Klinman for useful discussions. This work was supported by grants AI-12001, GM-19121 and HL-11119 from the NIH.

Received 29 June; accepted 9 October 1979.

- Klinman, N. R. & Press, J. L. *Transplant Rev.* **24**, 41–83 (1975).
- Nisonoff, A., Reichlin, M. & Margoliash, E. *J. biol. Chem.* **245**, 940–946 (1970).
- Urbanski, G. J. & Margoliash, E. *J. Immun.* **118**, 1170–1180 (1977).
- Eng, J. & Reichlin, M. *Molec. Immun.* **16**, 225–230 (1979).
- Jemmerson, R. & Margoliash, E. *J. biol. Chem.* (in the press).
- Reichlin, M., Nisonoff, A. & Margoliash, E. *J. biol. Chem.* **245**, 947–954 (1970).
- Nisonoff, A., Margoliash, E. & Reichlin, M. *Science* **155**, 1273–1275 (1967).
- Klinman, N. R. *J. exp. Med.* **133**, 963–972 (1971).
- Feldman, M. J. *exp. Med.* **136**, 532–545 (1972).
- Corradin, G. & Harbury, H. A. *Biochim. biophys. Acta* **221**, 489–496 (1970).
- Atassi, M. Z. *Immunochimistry* **12**, 423–438 (1975).
- Needleman, S. B. & Margoliash, E. *J. biol. Chem.* **241**, 853–863 (1966).

- Hennig, B. *Eur. J. Biochem.* **55**, 167–183 (1975).
- Carlson, S. S. *et al. Biochemistry* **16**, 1437–1442 (1977).
- Nolan, C. & Margoliash, E. *J. biol. Chem.* **241**, 1049–1059 (1966).
- Niece, R. L., Margoliash, E. & Fitch, W. M. *Biochemistry* **16**, 68–72 (1977).
- Nakashima, T., Higa, H., Matsubara, H., Benson, A. M. & Yasunobu, K. T. *J. biol. Chem.* **241**, 1166–1177 (1966).
- Margoliash, E., Smith, E. L., Kreil, G. & Tuppy, H. *Nature* **192**, 1125–1127 (1961).
- McDowall, M. A. & Smith, E. L. *J. biol. Chem.* **240**, 4635–4647 (1965).
- Hunter, W. M. & Greenwood, F. C. *Nature* **194**, 495–496 (1962).
- Kabat, E. A. & Wu, T. T. *Ann. N.Y. Acad. Sci.* **190**, 382–393 (1972).
- Dayhoff, M. O. & Eck, R. T. *Atlas of Protein Sequence and Structure* Vol. 5 (National Biomedical Research Foundation, Washington, DC, 1972).
- Borden, D. & Margoliash, E. in *Handbook of Biochemistry and Molecular Biology* (ed. Fasman, G. D.) 268–279 (Chemical Rubber Co., Cleveland, Ohio, 1976).

Complete nucleotide sequence of an influenza virus haemagglutinin gene from cloned DNA

A. G. Porter, C. Barber, N. H. Carey, R. A. Hallewell, G. Threlfall & J. S. Emtage

Searle Research Laboratories, Lane End Road, High Wycombe, Bucks, UK

A synthetic fowl plague virus (FPV) haemagglutinin gene has been cloned in bacteria and the complete sequence of the RNA gene deduced. It is 1,742 nucleotides long and the mRNA codes for 563 amino acids in an uninterrupted sequence. The nature of some of the important domains in the haemagglutinin has been established, and their structure is discussed in relation to their function. Extensive amino acid sequence homologies exist between FPV and human influenza haemagglutinins.

FOWL PLAGUE VIRUS (FPV) like other influenza viruses is a negative-strand virus with a genome of eight separate^{1–3} and unique RNA segments³, two of which code for the surface antigens, haemagglutinin (gene 4)^{4–6} and neuraminidase. The haemagglutinin (HA) is responsible for attaching the virus to cell receptors⁷ and for initiating infection^{8,9}. It is firmly established that host antibodies are directed against the haemagglutinin^{10,11}, and that the virus can escape neutralisation by spontaneous mutation, resulting in amino acid changes in the haemagglutinin. These changes may be major and cause pandemics (antigenic shift) or relatively minor and cause epidemics (antigenic drift)¹². The minor antigenic changes occur more frequently, but their precise nature and location remain unknown.

Due to the difficulty of RNA sequencing with a molecule the size of the HA gene, peptide sequencing was, until recently, the best method of studying antigenic variation of influenza viruses. Ward and Dopheide¹³ and Waterfield and Skehel¹⁴ have determined most of the peptide sequence of the mature HA of a Hong Kong strain (H3N2) and an Asian strain (H2N2), respectively. With the introduction of gene cloning and rapid DNA sequencing greater speed is obtained, and certain problems such as the difficulty of sequencing hydrophobic peptides are overcome. In addition, the gene sequence not only predicts the complete amino acid sequence of the mature protein, but will also indicate the sequence of any peptides that are discarded during its maturation and perhaps also if untranslated sequences (introns) interrupt the coding region of the gene.

Gene sequences can also provide information about the structure of genes not evident from the protein sequence, and a means of comparing different virus strains in greater detail than techniques such as hybridisation¹⁵, gel electrophoretic mobility¹⁶ and oligonucleotide mapping¹⁷.

In this article we describe the determination of the complete nucleotide sequence of the FPV haemagglutinin gene using recombinant DNA techniques. The predicted sequence of the HA allows us to correlate some of the known functions with specific structures in the molecule and to compare the amino acid sequence with partial sequences of human influenza haemagglutinins^{13,14}.

Gene 4 sequences in plasmid pBR322-FPV₄₋₁₀

We chose to polyadenylate the eight virion RNA segments of unfractionated FPV RNA, and reverse transcribe the mixture with an oligo(dT) primer, since the resultant transcripts included substantial numbers of full length or almost full length copies of all eight genes¹⁸.

Table 1 Use of codons in FPV gene 4

Phe UUU 13	Ser UCU 4	Tyr UAU 7	Cys UGU 7
UUC 13	UCC 4	UAC 8	UGC 9
Leu UUA 6	UCA 13	Ochre UAA 1	Opal UGA 0
UUG 9	UCG 1	Amber UAG 0	Trp UGG 8
Leu CUU 9	Pro CCU 3	His CAU 7	Arg CGU 1
CUC 4	CCC 1	CAC 4	CGC 0
CUA 4	CCA 8	Gln CAA 14	CGA 4
CUG 8	CCG 4	CAG 11	CGG 3
Ile AUU 17	Thr ACU 13	Asn AAU 23	Ser AGU 11
AUC 8	ACC 12	AAC 14	AGC 8
AUA 15	ACA 15	Lys AAA 22	Arg AGA 14
Met AUG 11	ACG 1	AAG 7	AGG 8
Val GUU 9	Ala GCU 12	Asp GAU 17	Gly GGU 5
GUC 6	GCC 3	GAC 9	GGC 10
GUA 5	GCA 17	Glu GAA 28	GGA 20
GUG 10	GCG 1	GAG 10	GGG 15

Codons with CG are underlined. 20.0% of codons used end in C, 32.9% end in A, 19.0% end in G and 28.1% end in U.

To clone individual genes, the S_1 -nuclease-treated double-stranded FPV cDNA was ligated to *Hind*III cut plasmid pBR322 DNA¹⁹ using the *Hind*III linker method²⁰ (see Fig. 1 legend). The mixture was then used to transform²¹ *Escherichia coli* K12 HB101. Thirteen tetracycline-sensitive colonies were obtained most of which contained plasmids with small gene inserts as judged by agarose gel electrophoresis. One plasmid, named pBR322-FPV₄₋₁₀, contained a large insert of about 1,700 nucleotides (Fig. 1, lane c) and the plasmid DNA specifically hybridised to a ³²P-labelled gene 4 probe²³ (see Fig. 1 legend for details). 59.7% of gene 4 probe was hybridised compared with not more than 3.1% for probes derived from genes 1-3 and 5-8, demonstrating the presence of gene 4

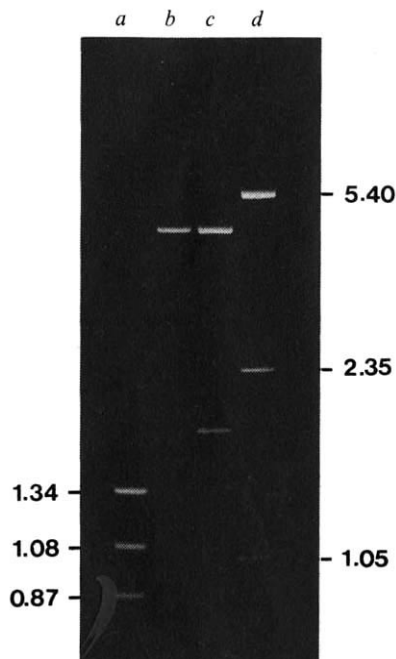


Fig. 1 Characterisation of the FPV gene 4 insert in pBR322 DNA. To prepare double-stranded cDNA from FPV RNA suitable for cloning, the total virion RNA from FPV (Rostock strain) was polyadenylated, and both cDNA strands were synthesised with reverse transcriptase, then treated with S_1 nuclease¹⁸. Approximately 0.2 µg cDNA was repaired in a 50-µl reaction containing: 50 mM Tris HCl pH 7.6, 10 mM MgCl₂, 50 µM of each deoxynucleoside triphosphate, 1 mM β-mercaptoethanol and 1 unit DNA polymerase I (Boehringer), for 10 min at 10°C, then ethanol precipitated. A synthetic *Hind*III linker (d[CCAAGCTTGG], Collaborative Research) was kinase-labelled and ligated to the gene mixture in a volume of 10 µl containing: 50 mM Tris HCl pH 7.6, 10 mM MgCl₂, 20 mM dithiothreitol (DTT), 1 mM ATP, approximately 0.2 µg dsDNA, 2 µM 5'-³²P-labelled linker and 1 µl T₄ DNA ligase (New England Biolabs) for 7 h at 15°C. After 7 and 23 h further linker was added to 2 µM, and at 23 h a further 1 µl ligase was added. Excess and unligated linkers were removed by *Hind*III digestion (30 units) at 37°C for 90 min, and chromatography through a G-150 column in 50 mM NaCl 0.2% SDS. The excluded peak containing genes was ethanol precipitated and used for ligation to pBR322 DNA as follows: 10 µg pBR322 DNA was digested with *Hind*III (25 units) and treated with bacterial alkaline phosphatase. After removing the enzymes with phenol, the linear dephosphorylated plasmid DNA (0.3 µg) was ligated to approximately 80 ng of gene mixture as above, but with 0.5 µl T₄ ligase in 20 µl for 16 h. 10 µl was withdrawn, and used to transform²¹ *E. coli* K12 HB101. The culture was plated on L-agar in the presence of 100 µg ml⁻¹ ampicillin. Plasmid DNA was purified from tetracycline-sensitive colonies and examined on a 1.4% agarose gel following *Hind*III restriction, an example of which is shown. a, *Hae*III cut ØX174 DNA markers in kilobase pairs; b, pBR322 + *Hind*III; c, pBR 322-FPV₄₋₁₀ + *Hind*III; d, *Hind*III cut PM2 DNA markers in kilobase pairs. To determine which FPV gene was inserted in each plasmid, ³²P-labelled gene probes were prepared and hybridised to plasmid DNA as follows: Plasmid DNA was first prepared by digesting 5 µg plasmid with *Eco*RI (12.5 units) and binding denatured 0.25 µg aliquots to small Millipore disks²³. 20 µg FPV RNA was electrophoresed in a 3% polyacrylamide gel, pH 3.5 in 7 M urea, and genes 1-3, 4, 5, 6, 7 and 8 eluted, incubated for 1 h at 0°C in 0.1 M NaOH, and labelled with ³²P at the newly exposed 5'-ends. Each ³²P-labelled probe (approximately 3,000 c.p.m.) was incubated with the DNA-coated Millipore disk at 37°C for 16 h in 100 µl 50% formamide in 3×SSC, 0.1% SDS and 2 µg tRNA. After extensive washing (50% formamide plus 5×SSC, 2×SSC, 2×SSC plus 10 µg ml⁻¹ pancreatic RNase and 2×SSC), the disks were dried and counted. Conditions for all restriction enzyme digestions were the same or very similar to those given in the catalogue of New England Biolabs Inc.

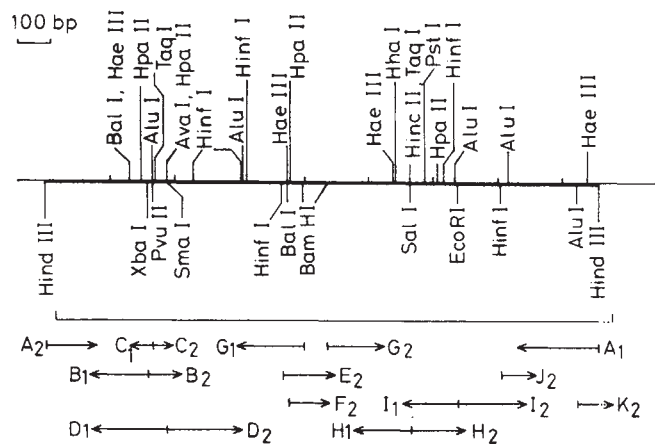


Fig. 2 Restriction enzyme map of the FPV gene 4 insert in pBR322. The insert is the thicker line, beneath which are shown the restriction sites labelled in the sequence analysis. The cloned gene 4 is 46 nucleotides shorter than full length gene 4 at the right hand end which corresponds to the 5'-terminal region of the virion RNA. The relative positions of the 'six base-pair' enzyme sites were first determined by standard mapping procedures. To locate sites for enzymes recognising 4 and 5 base pairs, the gene 4 insert (10 µg) was cut from the plasmid with *Hind*III (20 units) for 90 min at 37°C. After removal of enzyme, dephosphorylation and labelling of 5'-ends (Fig. 3), the insert was separated from the plasmid DNA by electrophoresis in a 1.4% agarose gel, and eluted²⁴. Subsequent digestion with, for example, *Hinf*I gave at least six fragments as shown by polyacrylamide gel electrophoresis (5% acrylamide, 0.17% bisacrylamide in 50 mM Tris-borate pH 8.3, 1 mM EDTA). The two *Hinf*I fragments containing the *Hind*III ends were identified from their ³²P label. The larger of these two was sensitive to *Sma*I, placing it at the left end of the insert. The positions of some of the other *Hinf*I sites were determined from the sensitivity of internal *Hinf*I fragments to *Sal*I and *Eco*RI. The distances sequenced are shown by arrows; those pointing to the left denote sequences derived from the virion RNA strand, and those pointing to the right from the cRNA strand. The vertical line interrupting each arrow indicates the restriction site used for labelling (see Fig. 3 legend for details). The dotted line in distance K₂ refers to the 5'-terminal 46 nucleotides of gene 4 missing from the plasmid. The line above the arrows indicates the position of full length gene 4 relative to the insert.

sequences in pBR322-FPV₄₋₁₀. This plasmid was chosen for sequence analysis as the insert was calculated¹⁸ to be an almost full length copy of gene 4.

Restriction enzyme mapping and sequence determination

As a first step to sequence analysis of pBR322-FPV₄₋₁₀, the cleavage sites for a number of restriction enzymes were mapped. Remarkably, all the six base-pair enzymes tested (except *Sac*I, *Hind*III and *Hpa*I) were found to cut the gene 4 insert at least once (Fig. 2). Some of these enzymes were used in the DNA sequence analysis (see Fig. 3 legend for details).

The lettered arrows in Fig. 2 refer to the distances sequenced and this procedure was usually done twice²². In all cases sequences from neighbouring restriction sites were overlapped by at least 15 nucleotides. It was useful to have the sequence on both strands at the overlap, where the bands on the sequencing gels (Fig. 3) were often close together. The few uncertainties were due either to compression²⁵ of bands at 'hairpin' loops or failure of a chemical reagent to cleave a phosphodiester bond²⁶ (at methylated C or A). These problems were overcome in all cases by sequencing the complementary DNA strand either in the opposite direction from a neighbouring restriction site, or in the same direction with 3'-end-labelled DNA²⁷ (see Fig. 4 legend for details).

We have confidence in the number of nucleotides present in the coding region because two protein synthesis reading frames are frequently blocked by nonsense triplets (Fig. 5), leaving one frame open to code for the protein sequence. Striking amino acid sequence homologies between FPV and human HA sequences^{13,14} also provide some indication of accuracy in parts of the DNA sequence (Fig. 5).

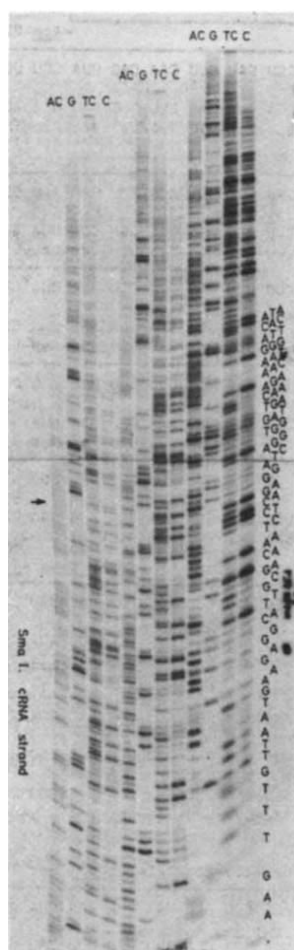
[illegible]

Fig. 4 Origin of the sequence at the right hand end of the gene 4 insert. *a*, The putative hairpin at the 5'-end of the first cDNA strand. *b*, The sequence found at the right hand end of the insert. The dotted lines in *a* refer to nucleotides that arose in the daughter molecules after the first round of plasmid replication. The result in *b* is two plasmids differing in positions 12-14, one with sequence *x* and the other sequence *y*. The complementary sequences for the right end of the insert going leftwards are shown in *c* (5' → 3') and *d* (3' → 5'). Evidence for heterogeneity is indicated only for the purines. Sequencing in the 3' → 5' direction (*d*) was done by 3'-end labelling of DNA³⁷. 20 µg of *Hind*III cut plasmid was heated at 90 °C for 5 min and rapidly cooled. The final incubation was: 0.4 mg ml⁻¹ DNA, 5 mM Tris, 0.02 mM EDTA, 0.1 M sodium cacodylate pH 6.9, 1 mM CoCl₂, 0.1 mM DTT, 100 µCi [α-³²P]CTP (>350 CiMole⁻¹) and 5 units deoxynucleotidyl terminal transferase for 1 h at 37 °C, then for 30 min at 37 °C with an additional 5 units of enzyme. Excess rC was removed with 5 µg pancreatic RNase plus 20 µg tRNA. The enzymes were removed, the DNA was recut with *Eco*RI (50 units) and the 0.447-kilobase fragment was purified by electrophoresis in a 1.4% agarose gel. *e*, The sequence determined from the RNA of gene 4 by priming with a 42-nucleotide restriction fragment derived from the gene 4 insert (see Fig. 3 legend). This sequence spans the right end of the gene 4 insert (*b*), and the five differences with the cloned DNA sequences are arrowed. The 42 base-pair *Alu*I (1,616) → *Hae*III (1,658) fragment was generated from the 5'-³²P-labelled *Hae*III (1,050) → *Hae*III (1,658) fragment by *Alu*I digestion (50 units) and purified by 5% polyacrylamide gel electrophoresis. Numbers refer to the distance in nucleotides of the restriction site from the 3'-end of the virionRNA strand (Fig. 5).] The fragment was then labelled at the 5'-ends (Fig. 3) and ethanol precipitated with 15-20 µg FVPV RNA. Denaturation was done at 80 °C for 2 min in 80% formamide, 0.01% SDS, 0.01M PIPES pH 6.5, 0.4 M NaCl, 1 mM EDTA and the mixture was annealed to the FVPV RNA at 50 °C for 20 min. Following ethanol precipitation, cDNA synthesis was carried out with reverse transcriptase¹⁸ and all four unlabelled deoxynucleoside triphosphates. To remove template RNA the cDNA/RNA was ethanol precipitated, briefly exposed to 0.3 M NaOH at 70 °C and neutralised at 20 °C. The large cDNA transcript was purified by 5% gel electrophoresis.

vRNA 3'-UCGUUUUCGUCCCCAUGUUU UAC UUG UGA GUU UAG GAC CAA AAG CGG GAA CAC CGU CAG UAG GGG UGU UUA CGU CUG UUU
 50
 cRNA 5'-AGCAAAAGCAGGGGUACAAA [AUG] AAC ACU CAA AUC CUG GUU UUC GCC CUU GUG GCA GUC AUC CCC ACA AAU GCA GAC AAA
 A/FPV Met-Asn-Thr-Gln-Ile-Leu-Val-Phe-Ala-Leu-Val-Ala-Val-Ile-Pro-Thr-Asn-Ala-Arg-Lys-(20)
 A/Japan Precursor peptide -Arg-Gln-

UAA ACA GAA CCU GUA GUA CGA CAU AGU UUA CCG UGG UUU CAU UUG UGU GAG UGA CUC UCU CCU CAU CAA CAG UUA CGU UGC
 100 150
 AUU UGU CUU GGA CAU CAU GCU GUA UCA AAU GGC ACC AAA GUA AAC ACA CUC ACU GAG AGA GGA GUA GAA GUU GUC AAU GCA ACG
 Ile-Cys-Leu-Gly-His-His-Ala-Val-Ser-Asn-Gly-Thr-Lys-Val-Asn-Thr-Leu-Thr-Glu-Arg-Gly-Val-Glu-Val-Val-Asn-Ala-Thr-(48)
 Ile-Cys-Ile-Gly-Tyr-His-Ala-Asx-Asn-

CUU UGU CAC CUC GCC UGU UUG UAG GGG UUU UAA ACG AGU UUU CCC UUU UCU UGG UGA CUA GAA CCG GUU ACG CCU GAC AAU CCC
 200 250
 GAA ACA GUG GAG CGG ACA AAC AUC CCC AAA AUU UGC UCA AAA GGG AAA AGA ACC ACU GAU CUU GGC CAA UGC GGA CUG UUA GGG
 Glu-Thr-Val-Gly-Arg-Thr-Asn-Ile-Pro-Lys-Ile-Cys-Ser-Lys-Gly-Lys-Arg-Thr-Thr-Asp-Leu-Gly-Gln-Cys-Gly-Leu-Leu-Gly-(76)
 UGG UAA UGG CCU GGU GGA GUU ACG CUG GUU AAA GAU CUU AAA AGU CGA CUA GAU UAU UAG CUC UCU GCU CUU CCU UUA CUA CAA
 HpaII XbaI PvuII 300 TspI
 ACC AAU ACC GGA CCA CCU CAA UGC GAC CAA UUU CUA GAA UUU UCA GCU GAU CUA AUA AUC GAG AGA CGA GAA GGA AAU GAU GUU
 Thr-Ile-Thr-Gly-Pro-Gln-Cys-Asp-Gln-Phe-Leu-Glu-Phe-Ser-Ala-Asp-Leu-Ile-Ile-Glu-Arg-Arg-Glu-Gly-Asn-Asp-Val-(104)

ACA AUG GGC CCC UUC AAA CAA UUA CUU CUC CGU AAC GCU GUU UAG GAG UCU CCU AGU CCA CCC UAA CUG UUU CUU UGU UAC CCU
 350 400
 UGU UAC CCG GGG AAG UUU GUU AAU GAA GAG GCA UUG CGA CAA AUC CUC AGA GGA UCA GGU GGG AUU GAC AAA GAA ACA AUG GGA
 Cys-Tyr-Pro-Gly-Lys-Phe-Val-Asn-Glu-Glu-Ala-Leu-Arg-Gln-Ile-Leu-Arg-Gln-Gly-Ser-Gly-Ile-Asp-Lys-Glu-Thr-Met-Gly-(132)
 AAG UGU AUA UCA CCU UAU UCC UGG UUG UGU UGA UCA CGU ACA UCU AGU CCC AGA AGU AAG AUA CGU CUU UAC CUC ACC
 450 500
 UUC ACA UAU AGU GGA AUA AGG ACC AAC GGA ACA ACU AGU GCA UGU AGA AGA UCA GGG UCU UCA UUC UAU GCA GAA AUG GAG UGG
 Phe-Thr-Tyr-Ser-Gly-Ile-Arg-Thr-Asn-Gly-Thr-Ser-Ala-Cys-Arg-Arg-Ser-Gly-Ser-Ser-Phe-Tyr-Ala-Glu-Met-Glu-Trp-(160)
 -Asx-Met-Val-Trp-

GAG GAC AGU UUA UGU CUG UUA CGA AGA AAG GGU GUU UAC UGU UUU AGU AUG UUU UUG UGU UCC UCU CUU AGU CGA GAC UAU CAG
 550
 CUC CUG UCA AAU ACA GAC AAU GCU UCU UUC CCA CAA AUG ACA AAA UCA UAC AAA AAC ACA AGG AGA GAA UCA GCU CUG AUA GUC
 Leu-Leu-Ser-Asn-Thr-Asp-Asn-Ala-Ser-Phe-Pro-Gln-Met-Thr-Lys-Ser-Tyr-Lys-Asn-Thr-Arg-Arg-Glu-Ser-Ala-Leu-Ile-Val-(188)
 Leu-Thr-Lys-
 -Gly-Ser-Tyr-Thr-Asn(Asn, Ser, Gly, Glx)Met -Leu-Ile-Ile-
 A/Memphis -Met-Pro-Asn-Asn-Asp-Asn-Phe-Asp-Lys-Leu-Tyr-Ile-

ACC CCU UAG GUG GUA AGU CCU AGU UGG UGG - CUU GUC UGG UUU GAU AUA CCC UCA CCU UUA UUU GAC UAU UGU CAG CCC UCA
 600 650
 UGG GGA AUC CAC CAU UCA GGA UCA ACC ACC - GAA CAG ACC AAA CUA UAU GGG AGU GGA AAU AAA CUG AUA ACA GUC GGG AGU
 Trp-Gly-Ile-His-His-Ser-Gly-Ser-Thr-Thr - Glu-Gln-Thr-Lys-Leu-Tyr-Gly-Ser-Gly-Asn-Lys-Leu-Ile-Thr-Val-Gly-Ser-(215)
 Trp-Gly-Val-His-His-Pro-Ile-Asp-Glu-Thr - Glu-Gln-Arg-
 Trp-Gly-Val-His-His-Pro - Ser-Thr-Asp-Gln-Glu-Gln-Thr-Ser-Leu-Tyr-Val-Gln-Ala-Ser-Gly-Arg-Val-Thr-Val-Ser-Thr-

AGG UUU AUA GUA GUU AGA AAA CAC GGC UCA GGU CCU UGU GCU GGC GUC UAU UUA CCG GUC AGG CCU GCC UAA CUA AAA GUA ACC
 700 750
 UCC AAA UAU CAU CAA UCU UUU GUG CCG AGU CCA GGA ACA CGA CCG CAG AUA AAU GGC CAG UCC GGA CCG AUU GAU UUU CAU UGG
 Ser-Lys-Tyr-His-Gln-Ser-Phe-Val-Pro-Ser-Pro-Gly-Thr-Arg-Pro-Gln-Ile-Asn-Gly-Gln-Ser-Gly-Arg-Ile-Asp-Phe-His-Trp-(243)
 Met-Gln-Phe-Ser-Trp-
 Lys-Arg-Ser-Gln-Gln-Thr-Ile-Ile-Pro-Asn-Ile-Gly-Ser-Arg-Pro-Trp-Val-Arg-Gly-Leu-Ser-Ser-Arg-Ile-Ser-Ile-Tyr-Trp-

AAC UAG AAC CUA GGG UUA CUA UGU CAA UGA AAA UCA AAG UUA CCC CGA AAG UAU CGA GGU UUA GCA CCG UCG AAG AAC -
 800
 UUG AUC UUG GAU CCC AAU GAU ACA GUU ACU UUU AGU UUC AAU GGG GCU UUC AUA GCU CCA AAU CGU GCC AGC UUC UUG -
 Leu-Ile-Leu-Asp-Pro-Asn-Asp-Thr-Val-Thr-Phe-Ser-Phe-Asn-Gly-Ile-Ala-Phe-Asn-Arg-Ala-Ser-Phe-Leu - (269)
 Thr-Leu-Leu-Asp-Met-Trp-Asp-Thr-Ile-Asn-Phe-Glu-Ser-Thr-Gly-Asn-Leu-Ile-Ala-Pro-Glu - Tyr-Phe-Lys-Ile-Ser-Lys-
 Thr-Ile-Val-Ile-Pro-Gly-Asp-Ile-Leu-Val-Ile-Asn-Ser-Asx-Gly-Asx-Leu-Ile-Ala-Pro-Arg-Gly-Tyr-Phe-Lys-Met -

UCC - CCU UUC AGG UAC CCC UAG GUC - UCG CUA CAC GUC CAA CUA CGA UUA ACG CUU CCC CUU ACG AUG GUG UCA CCU CCC
 850 900
 AGG - GGA AAG UCC AUG GGG AUC CAG - AGC GAU GUG CAG GUU GAU GCU AAU UGC GAA GGG GAA UGC UAC CAC AGU GGA GGG
 Arg - Gly-Lys-Ser-Met-Gly-Ile-Gln - Ser-Asp-Val-Gln-Val-Asp-Ala-Asn-Cys-Glu-Gly-Glu-Cys-Tyr-His-Ser-Gly-Gly-(295)
 Arg - Gly - Ser-Ser-Gly-Ile-Met-Lys - Thr-Glu-Gly-Thr-Leu-Glu-Asn-Cys-Glu-Thr-Lys-Cys-Glu-Thr-Pro-Leu-Gly-
 Pro-Thr-Gly-Lys-Ser-Ser - Ile-Met-Arg-Ser-Asp-Ala - Pro-Ile-Gly-Thr-Cys-Ile-Ser-Glu-Cys-Ile-Thr-Pro-Asn-Gly-

UGA UAU UGU UCG UCU AAC GGA AAA GUU UUG UAU UUA UCG UCU CGU CAA CCG UUU ACG GGU UCU AUA CAU UUU GUC CUU UCA AAU
 950
 ACU AUA ACA AGC AGA UUG CCU UUU CAA AAC AUA AAU AGC AGA GCA GUU GGC AAA UGC CCA AGA UAU GUA AAA CAG GAA AGU UUA
 Thr-Ile-Thr-Ser-Arg-Leu-Pro-Phe-Gln-Asn-Ile-Asn-Ser-Arg-Ala-Val-Gly-Lys-Cys-Pro-Arg-Tyr-Val-Lys-Gln-Glu-Ser-Leu-(323)
 Ala-Ile-Asn-Thr-Thr-Leu-Pro-Phe-His-Asx-Val-His-Pro-Leu-Thr-Ile-Gly-Glx-Cys-Pro-Lys-Tyr-Ser - Glu-Lys - Leu-
 Ser-Ile-Pro-Lys-Pro-Asp-Asp-Phe-Gln-Asn-Val-Asn-Lys-Ile-Thr-Tyr-Gly-Ala-Cys-Pro-Lys-Tyr-Val-Lys-Gln-Asn-Thr-Leu-

Evidence for repair and heterogeneity in cloned DNA

The first sequences to be established for the gene 4 insert were at the extremities (Fig. 2). The left end of the insert (adjacent to the *EcoRI* site of pBR322 DNA) was found to have the sequence corresponding exactly to the 3'-end of the virionRNA (5'-end of cRNA)²⁸, but the right hand end could not be matched with the published sequence at the 5'-end of the virionRNA²⁸. Therefore an experiment was devised to determine how much of gene 4

was missing. A 42 base-pair restriction fragment from between the *AluI* and *HaeIII* sites nearest the right end of the insert (Fig. 2) was purified, labelled and used as a primer for reverse transcriptase in the presence of FPV RNA and all four unlabelled deoxynucleoside triphosphates¹⁸ (see Fig. 4 legend for details). The strand corresponding to the cRNA presumably annealed to the RNA of gene 4 since the DNA product appeared as a single band in a 5% polyacrylamide gel (not shown). Its length was estimated to be about 140 nucleotides. The band was eluted and the DNA sequenced²² (Fig. 4e). The

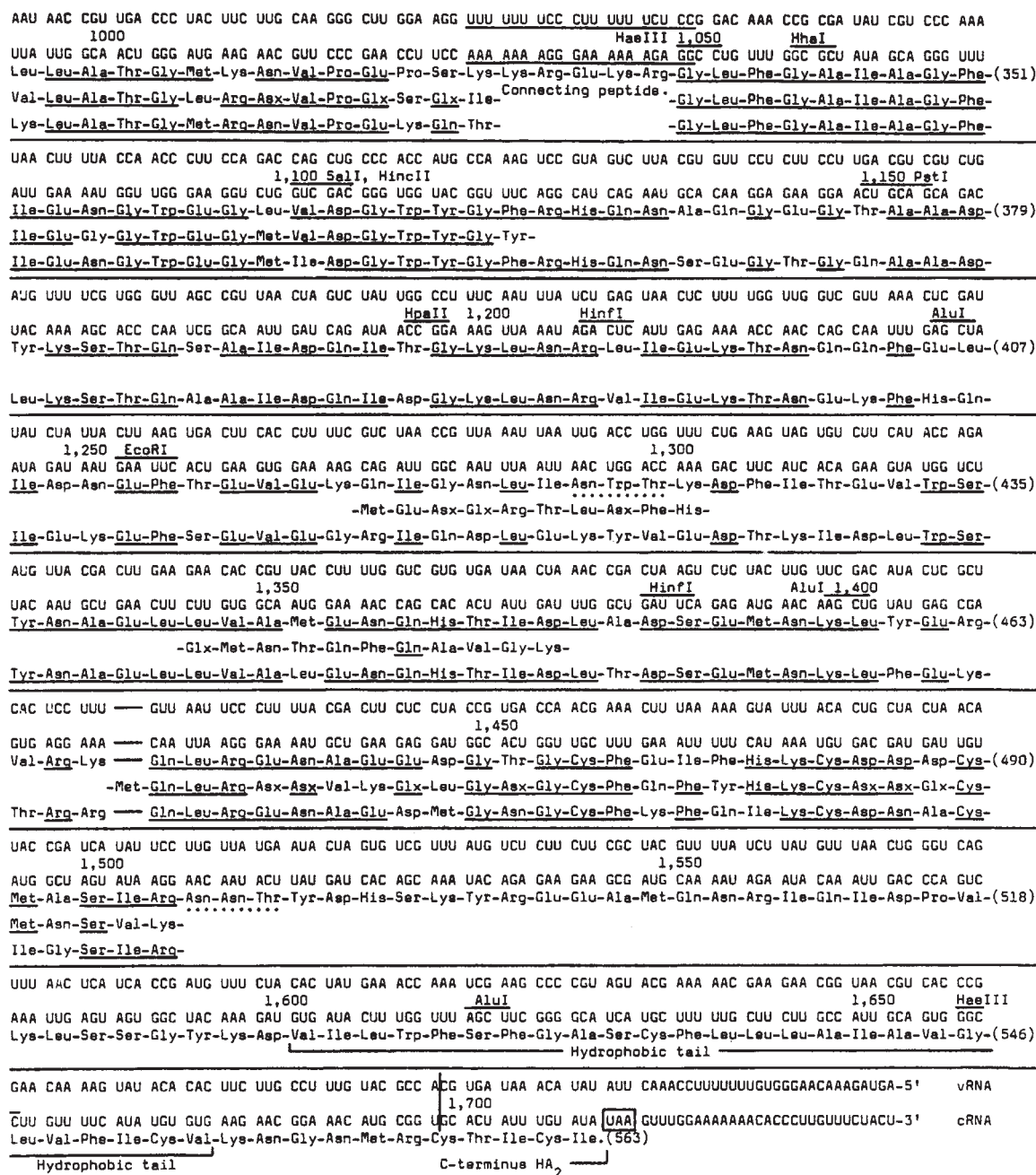


Fig. 5 Nucleotide sequence of the FPV HA gene and comparison of the predicted amino acid sequence of the HA with human HA sequences. The virus RNA (with the 3'-end to the left) is shown above the complementary RNA. Boxes show the initiation codon (AUG) and termination codon (UAA). The positions of the restriction enzyme sites (Fig. 2) found in the corresponding plasmid DNA are indicated by a bar above the cRNA. The heavy underlining near the initiation codon shows two areas of homology with prokaryotic ribosome binding sites. The underlined purine-rich sequence in the cRNA (nucleotides 1,029–1,049) spans the region coding for the peptide connecting HA₁ to HA₂. The vertical line to the left of the termination codon shows the position corresponding to the attachment of the *Hind*III linker to the incomplete DNA copy of gene 4. To the left of this line is shown the actual sequence of the RNA gene (Fig. 4e). The predicted amino acid sequence of the immature FPV HA is immediately below the cRNA. The A/Japan/305/57 (H2N2) partial sequence is shown below the complete FPV sequence, and the third line has the A/Memphis/102/72 (H3N2) partial sequence. In maximising homology (underlining) the sequences of these HAs have been aligned according to the positions of the cysteines. Dotted lines below the FPV HA sequence indicate asparagines which may be glycosylated in the mature HA.

23 nucleotides at its 3'-terminal end are complementary to the published sequence of 23 nucleotides at the 5'-end of FPV gene 4, demonstrating that the strong transcription stop had occurred at the 5'-end of the template RNA²⁸. The results show that the FPV gene 4 insert is only 46 nucleotides shorter than the full length gene (Fig. 5). The sequence of 70 nucleotides at the 5'-terminus derived in this way is identical to the sequence determined for the 5'-end of gene 4 by rapid RNA sequencing⁵³. However, five differences were found between the sequence deduced by these two independent methods (Fig. 4e) and the

sequence near the right hand end of the insert determined on both cloned DNA strands (Fig. 4c, d). These differences (arrowed in Fig. 4e) are from top to bottom: absence of a G·C pair in the cloned DNA (Fig. 4c, d, between positions 3 and 4); an A instead of a T in the cloned DNA (Fig. 4d, position 5); and sequence heterogeneity in three positions in the cloned DNA. Both an A and a G are seen at position 13 (Fig. 4c) and at positions 12 and 14 (Fig. 4d), whereas elsewhere on these sequencing gels cuts are made only at A or G. Heterogeneity in the corresponding pyrimidines is not so obvious since cuts after

C are normally made in both pyrimidine-specific reactions²².

It is extremely unlikely that these base changes resulted from reverse transcriptase errors, because the conditions of reverse transcription of FPV RNA using the 42-base-pair primer were identical to those used in synthesis of double-stranded DNA for cloning. We suggest instead the following model to explain why the base changes occurred. As with other cDNAs, a hairpin will form and prime the synthesis of the second cDNA strand²⁹. The structure in Fig. 4a is the only one that can account for the loss of 46 nucleotides from the 5'-end of the gene, provided that the following assumptions are made. First, in the molecule which was cloned the three single-stranded 'bubbles', but not the large terminal loop, survived S₁ nuclease treatment, and second, reverse transcriptase stabilised the interaction of the 3'-terminal T with an A on the opposite strand so that synthesis of the second cDNA strand could commence.

In the proposed hairpin structure (Fig. 4a) there are three regions where mismatching occurs, and these regions coincide exactly with the five base changes found in the cloned DNA. The largest 'bubble' (positions 12–14) corresponds to the three positions where sequence heterogeneity is observed (Fig. 4c, d). The heterogeneity reflects the presence of two plasmids in the bacteria, and it seems probable that these would have arisen after the first round of DNA replication, since it is highly unlikely that one bacterial cell would have been transformed by two plasmid molecules of identical lengths derived from a heterogeneous mixture of reverse transcripts¹⁸. Thus, the mismatched AC pairs at positions 13 and 14 (Fig. 4a) would segregate to a TA pair in one new DNA molecule and a CG pair in the other. Likewise, the GT pair at position 12 becomes a GC in one new DNA molecule and a TA in the other.

In the other two regions excisions could have taken place to restore a perfect duplex (Fig. 4b). Thus, the extra C residue nearest the S₁ cleavage site is excised, and two residues to the left (position 5) a T in the strand corresponding to the cRNA is replaced with an A, which can now base-pair with the T on the opposite DNA strand. In contrast to positions 12–14, there is no evidence of sequence heterogeneity, suggesting that the excisions occurred before replication.

The nucleotide sequence of FPV gene 4

The FPV gene 4 is 1,742 nucleotides long (Fig. 5). This is about 300 nucleotides shorter than a previous estimate¹⁸ of 2,060 nucleotides based on migration rate of RNA in formamide–polyacrylamide gels. The coding region begins with an AUG and is terminated by a single nonsense codon (UAA). Around nucleotides 1,030–1,050 there is an unusual sequence in the mRNA—an uninterrupted stretch of purines most of which code for the peptide connecting the two HA subunits, HA₁ and HA₂ (ref. 13).

During *in vivo* transcription of viral RNA, host mRNAs prime the synthesis of influenza mRNAs adding an extra 10–15 nucleotides to each 5'-end³⁰ (not shown). At the other end, transcription terminates prematurely³¹, probably at the U₇ sequence 17–23 nucleotides from the 5'-end of the virionRNA template. Thus the 5'- and 3'-non-coding regions (including the UAA) are 31–36 and approximately 9 nucleotides long, respectively. The length of the 3' region is therefore short compared with the corresponding region of other mRNAs found in eukaryotic cells^{26,29,32}. A putative signal sequence 5'AAU-AAA3' has been found in the 3'-non-coding region of many different types of eukaryotic mRNA^{26,29,32}. It is absent from FPV gene 4 (Fig. 5) and this resembles the situation in some picornaviruses³³ and one of the mRNAs of vesicular stomatitis virus³⁴.

The 3'-ends of the cRNA and virionRNA of gene 4 can be aligned so that 12 out of 15 nucleotides are the same. This is significant as a possible signal for the initiation of replication on both these RNAs as has already been noted^{28,31}. Homology between 3'-ends also means that the 5'- and 3'-ends of either virionRNA or cRNA are complementary²⁸, but there is no known function for such an interaction.

A striking feature of the 5'-non-coding region near the AUG (underlined in Fig. 5) is the homology with some prokaryotic ribosome binding sites. These include the R17 replicase³⁵, T₇ early mRNA³⁶ and *E. coli* Trp E sites³⁷. In FPV gene 4, there is a characteristic purine-rich sequence, AGGGG, which interacts with the 3'-end of 16S rRNA in prokaryotes³⁸, followed immediately by UUACA. The latter sequence is present in the prokaryotic sites as UUAC and U-ACA. There is even greater homology between the gene 4 site and a spurious ribosome binding site in Q β RNA which lacks an AUG³⁹. Here at least 12 out of 16 nucleotides are the same. Support for the idea that the gene 4 site is prokaryotic in character comes from experiments on the expression of FPV gene 4 in *E. coli*⁴⁰.

Codon usage

The choice between degenerate codons in many mRNAs is frequently non-random, with markedly different preferences in different mRNAs^{26,29,32}. In some mRNAs, the most frequently used codons correlate with the abundant species of isoaccepting tRNA⁴¹, while in others the usage reflects the base composition⁴². FPV gene 4 is U-rich (34.3%) and has a low G content (18.5%). Reflecting this, there is a significant bias for codons with A in the third position (notably GAA, AGA and AAA), and against codons with third position C (Table 1).

Gene 4 has an overall deficiency of CG relative to the other 15 dinucleotides. This is not due only to the low G content, since GG sequences are almost twice as abundant. Again reflecting the base composition of the virionRNA, codons with CG in first and second or second and third position rarely occur (Table 1). There is also a deficiency of CG in the eukaryotic genome⁴³. In this case it has been suggested that the frequency of CG has been kept to a minimum since cytosine in CG is a major site of DNA methylation and methylated nucleotides are 'hot spots' for mutation³². But this cannot be true for influenza, an RNA virus with no known DNA intermediate in its replication.

Amino acid sequence organisation of the FPV haemagglutinin

The immature FPV HA consists of 563 amino acids, of which 18 form the N-terminal 'pre-peptide', 319 are in HA₁, 5 are in a connecting peptide and 221 are in HA₂ (Fig. 5). The amino acid composition shows similarities to human haemagglutinins; for example in the high proline content of HA₁ compared to HA₂, and in the number and distribution of cysteine residues^{13,14}. The positions of the cysteines have been highly conserved, perhaps to help give each molecule a basically similar overall shape by disulphide bridging.

Existing evidence suggests that immature HA has a short N-terminal leader sequence which is cleaved off when the process of membrane insertion is complete¹⁴. The DNA sequence predicts that the FPV leader consists of 18 amino acids, none of which is polar (Fig. 5). From the fifth residue, there is a core of 11 consecutive hydrophobic amino acids. The leader sequence of another influenza A haemagglutinin (RI/57/57) is 15 residues in length and shows almost no sequence homology with the FPV leader, but like FPV has an uninterrupted sequence of hydrophobic amino acids⁴⁴. The size and non-polar character of these N-terminal peptides is very similar in other secreted proteins⁴⁵, which lends strong support to the signal peptide hypothesis⁴⁶. This states that the process of membrane transport is initiated by insertion of a hydrophobic pre-peptide, which is subsequently cleaved off.

Cleavage of HA into HA₁ and HA₂ usually takes place either on smooth internal membranes or at the plasma membrane⁴⁷. This step is essential for infectivity⁹ and apparently also for pathogenicity⁴⁸. It is carried out by an unknown host protease with a similar specificity to trypsin⁴⁷. In FPV, the sequence around the cleavage site is Lys-Lys-Arg-Glu-Lys-Arg-Gly (Fig. 5), where the first residue becomes the C-terminus of HA₁ and the last residue the N-terminus of HA₂. The basic nature of these residues accounts for the *in vitro* trypsin cleavage of HA

into HA₁ and HA₂, but it is not known whether only some or all the basic residues are cut either *in vitro* or *in vivo*. In contrast to FPV, the C-terminal amino acid of HA₁ of A/Japan/305/57 is isoleucine¹⁴ and A/Memphis/102/72 has C-terminal threonine¹³. These differences could explain why the susceptibility of HA to cleavage depends on the virus strain⁴⁷ as well as on the cell type⁴⁸.

In human influenza A strains, the C-terminal part of HA₂ has not yet been sequenced owing to the insolubility of the peptides derived from it, indicating the presence of an extensive hydrophobic region^{13,14}. In FPV this begins at residue 527 and extends for 26 amino acids to residue 552 (Fig. 5). Of these, 20 residues are hydrophobic, and are probably involved in anchoring the HA to the viral lipoprotein envelope¹⁴.

There are five potential sites for carbohydrate attachment in HA₁ and two in HA₂ (Fig. 5). This conclusion is based on the presence of the sequences Asn-X-Ser and Asn-X-Thr, which are known to occur in other glycoproteins (including haemagglutinins^{13,14}) at sites of N-glycosidic linkage of asparagine to oligosaccharides¹⁴. Previous studies on the HA from FPV (Dutch) have indicated that there are five or six separate oligosaccharide units, but their exact location was not determined⁴⁹. In A/Japan/305/57, the approximate location of carbohydrate side chains is known to be residues 11, 23, 192 and 303 in mature HA₁ and very near the bromelain cleavage site in HA₂ (ref. 14). In FPV, the sites at residues 30 and 46 in HA₁, and residue 496 in HA₂ seem to correspond to the known A/Japan sites at 11 and 23 in mature HA₁ and the single site in HA₂, respectively.

In Fig. 5, the FPV HA sequence is compared with published partial HA sequences for an H2N2 strain (A/Japan/305/57) and an H3N2 strain^{13,14} (A/Memphis/102/72). There is considerable sequence homology (38%) between all strains. This occurs to an even greater extent between FPV and Japan (48%), Memphis and Japan (50%) and FPV and Memphis (55.5%). Conservation of sequence appears to be greater in HA₂ than HA₁, which is not surprising as only HA₁ is thought to undergo antigenic variation.

As pointed out previously, the N-terminal 24 residues of several HA₂ molecules are highly conserved¹⁴, and this now extends to FPV. The cleavage into HA₁ and HA₂ may expose the hydrophobic sequence at the N-terminus of HA₂, which then enables the HA to penetrate the cell membrane¹⁴. In at least 63% of positions where there is an amino acid difference between any two HAs (Fig. 5), the change can be accounted for by a single nucleotide substitution. This is significantly higher

than the value of 38% for all possible pairs of amino acids. This consideration, together with the overlapping nature of the homologies between the three HAs (Fig. 5), implies that the corresponding HA genes have evolved from one and not from two or more distinct ancestors.

Concluding remarks

Gene 4 of FPV is the first influenza virus gene to be completely sequenced and it predicts for the first time the complete amino acid sequence of a viral HA. The gene has some interesting features. The most notable of these are seen in the mRNA—a short 3' non-coding region, a prokaryotic-like ribosome binding site and an apparent lack of introns.

The sequence analysis of one end of the cloned gene 4 shows clearly that base changes and a base deletion occurred during the cloning process. The model to account for these changes is compelling, implying that because the changes are related to the hairpin at the 5'-end of the first cDNA strand they will not be found elsewhere in the synthetic gene. The results in any event indicate the importance of sequencing RNA independently of cloned DNA.

The HA of influenza viruses is a remarkable multifunctional protein. Our results and those of others have built up a picture of several functionally important domains in the molecule. These are the three hydrophobic regions with distinct functions, the connecting peptide and the carbohydrate attachment sites.

The region of HA that interacts with cell receptors at an early stage in infection remains to be determined, as do the antigenic sites, now known to be in the N-terminal half of HA₁ in a Hong Kong strain⁵⁰. The FPV sequence does, however, provide a framework for determining these sites in combination with other approaches. For example, it will now be possible to determine the primary structure of HA genes from influenza strains that differ only in their reactivity towards monoclonal antibodies⁵¹. A single nucleotide change would show which amino acid had changed and pinpoint the antigenic site. Ultimately, these methods may enable us to predict which amino acid changes will produce influenza viruses capable of causing epidemics.

We thank Mr G. Catlin and Mr B. Jenkins for technical help, Dr M. Waterfield for useful discussions, Drs G. Air and J. Robertson for allowing us access to their unpublished results, Dr J. Robertson for FPV RNA and Dr A. J. Hale for providing research facilities. All cloning experiments were done in a category III laboratory inspected and approved by GMAG. This report has been published elsewhere⁵².

Received 16 August; accepted 14 September 1979.

- Inglis, S. C., Carroll, A. R., Lamb, R. A. & Mahy, B. W. J. *Virology* **74**, 489–503 (1976).
- Scholtissek, C. *Curr. Topics Microbiol. Immun.* **80**, 139–169 (1978).
- McGeoch, D., Fellner, P. & Newton, C. *Proc. natn. acad. Sci. U.S.A.* **73**, 3045–3049 (1976).
- Scholtissek, C., Harms, E., Rhode, W., Orlich, M. & Rott, R. *Virology* **74**, 332–344 (1976).
- Inglis, S. C., McGeoch, D. & Mahy, B. W. J. *Virology* **78**, 522–536 (1977).
- Almond, J. W. & Barry, R. D. *Virology* **92**, 407–415 (1979).
- Hirst, G. K. *J. exp. Med.* **75**, 49–64 (1942).
- Klenk, H. D., Rott, R., Orlich, M. & Blödmann, J. *Virology* **68**, 426–439 (1975).
- Lazarowitz, S. G. & Choppin, P. W. *Virology* **68**, 440–454 (1975).
- Drzeniek, R., Seto, J. T. & Rott, R. *Biochim. biophys. Acta* **128**, 547–558 (1966).
- Laver, W. G. & Kilbourne, E. D. *Virology* **30**, 493–501 (1966).
- Laver, W. G. & Webster, R. G. *Br. med. Bull.* **35**, 29–33 (1979).
- Ward, C. W. & Dopheide, T. A. *Br. med. Bull.* **35**, 51–56 (1979).
- Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. *Br. med. Bull.* **35**, 57–63 (1979).
- Scholtissek, C., Rohde, W., Von Hoyningen, V. & Rott, R. *Virology* **87**, 13–20 (1978).
- Palese, P. & Schulman, J. L. *J. Virol.* **17**, 876–884 (1976).
- Nakajima, K., Desselberger, U. & Palese, P. *Nature* **274**, 334–339 (1978).
- Emtage, J. S., Catlin, G. H. & Carey, N. H. *Nucleic Acids Res.* **6**, 1221–1240 (1979).
- Boyer, H. W. *et al. in Recombinant Molecules: Impact on Science and Society* (eds Beers, R. F. & Bassett, E. G.) 9–20 (Raven, New York, 1977).
- Scheller, R. H., Dickerson, R. E., Boyer, H. W., Riggs, A. D. & Itakura, K. *Science* **196**, 177–180 (1977).
- Hallewell, R. A. & Emtage, J. S. *Gene* (in the press).
- Maxam, A. M. & Gilbert, W. *Proc. natn. acad. Sci. U.S.A.* **74**, 560–564 (1977).
- Gillespie, D. & Spiegelman, S. *J. molec. Biol.* **12**, 829–842 (1965).
- Vogelstein, B. & Gillespie, D. *Proc. natn. acad. Sci. U.S.A.* **76**, 615–619 (1979).
- Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107–110 (1978).
- McReynolds, L. *et al. Nature* **273**, 723–728 (1978).
- Roychoudhury, R., Jay, E. & Wu, R. *Nucleic Acids Res.* **3**, 863–877 (1976).
- Skehel, J. J. & Hay, A. J. *Nucleic Acids Res.* **5**, 1207–1219 (1978).
- Seeborg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 486–494 (1977).
- Plotch, S. J., Bouloy, M. & Krug, R. M. *Proc. natn. acad. Sci. U.S.A.* **76**, 1618–1622 (1979).
- Skehel, J. J. & Hay, A. J. *J. gen. Virol.* **39**, 1–8 (1978).
- Heindell, H. C., Liu, A., Paddock, G. V., Studnicka, G. M. & Salser, W. A. *Cell* **15**, 43–54 (1978).
- Porter, A. G. *et al. Nature* **276**, 298–301 (1978).
- McGeoch, D. J. & Turnbull, N. T. *Nucleic Acids Res.* **5**, 4007–4024 (1978).
- Steitz, J. A. *Nature* **224**, 957–964 (1969).
- Arrand, J. R. & Hindley, J. *Nature new Biol.* **244**, 10–13 (1973).
- Squires, C. J. *molec. Biol.* **103**, 351–381 (1976).
- Shine, J. & Dalgarno, L. *Nature* **254**, 34–38 (1975).
- Steitz, J. A. *J. molec. Biol.* **73**, 1–16 (1973).
- Emtage, J. S. *et al. Nature* (in the press).
- Post, L. E., Strycharz, G. D., Nomura, M., Lewis, H. & Dennis, P. P. *Proc. natn. acad. Sci. U.S.A.* **76**, 1697–1701 (1979).
- Nakanishi, S. *Nature* **278**, 423–427 (1979).
- Subak-Sharpe, J. H. *Br. med. Bull.* **23**, 161–168 (1967).
- Air, G. *Virology* **97**, 468–472 (1979).
- Strauss, A. W., Bennett, C. D., Donohue, A. M., Rodkey, J. A. & Alberts, A. W. *J. biol. Chem.* **252**, 6846–6855 (1977).
- Blobel, G. & Dobberstein, B. *J. cell Biol.* **67**, 835–851 (1975).
- Rott, R., Klenk, H. D. & Scholtissek, C. in *The Influenza Virus Hemagglutinin* (eds Laver, W. G., Bachmayer, H. & Weil, R.) 69–81 (Springer, New York, 1978).
- Bosch, F. X., Orlich, M., Klenk, H. D. & Rott, R. *Virology* **95**, 197–207 (1979).
- Klenk, H. D., Schwarz, R. T., Schmidt, M. F. G. & Wollert, W. in *The Influenza Virus Hemagglutinin* (eds Laver, W. G., Bachmayer, H. & Weil, R.) 83–99 (Springer, New York, 1978).
- Jackson, D. C., Dopheide, T. A., Russell, R. J., White, D. O. & Ward, C. W. *Virology* **93**, 458–465 (1979).
- Gerhard, W. in *The Influenza Virus Hemagglutinin* (eds Laver, W. G., Bachmayer, H. & Weil, R.) 15–24 (Springer, New York, 1978).
- Porter, A. & Emtage, J. S. *Biochem. Soc. Bull.* **1**, 72 (1979); *Nucleotide Group Mng.* Cambridge (1979).
- Robertson, J. S. *Nucleic Acids Res.* **6**, 3745–3757 (1979).

Transposable mating type genes in *Saccharomyces cerevisiae*

James Hicks, Jeffrey N. Strathern & Amar J. S. Klar

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

*A functional copy of the α mating type gene of *Saccharomyces cerevisiae* has been cloned by transformation in yeast. Using the Southern blotting procedure it has been shown that three distinct genetic loci implicated in mating type interconversion (HML, HMR and MAT) contain sequences homologous to the cloned fragment. The restriction fragment associated with each locus exhibits a characteristic size which can be correlated with the mating type allele present at that locus. The characteristic size difference between the a and α genetic elements made it possible to demonstrate that the homothallic interconversion of mating types in this yeast occurs by DNA rearrangement as proposed in the 'cassette hypothesis'.*

In the life cycle of homothallic baker's yeast a simple example of cell differentiation occurs in which haploid cells of one mating type (a or α) give rise to progeny of the opposite type during mitotic growth. Subsequent conjugation between siblings of opposite cell type establishes the a/α diploid phase. Genetic observations on this regulatory switching process have led to the proposal that the mating type genes reside on translocatable DNA elements and that switches between cell types are accomplished by transposition of these elements^{1,2}. Although a great deal of genetic evidence has been obtained in support of this 'cassette model' for mating type interconversion, direct demonstration of the proposed transposition mechanism requires identification and cloning of the DNA sequences involved. Using recently developed methods for yeast transformation, we have isolated a copy of the α mating type gene and have used this clone as a hybridisation probe to analyse the arrangement of the mating type genes in the yeast genome and to document that genetic rearrangements are involved in the mating type interconversion process. Our analysis is consistent with the prediction of the cassette model that silent copies of the mating type genes reside at the *HML* and *HMR* loci on chromosome III and that replicas of these unexpressed sequences are translocated to the mating type locus when switching occurs.

Background of the cassette hypothesis

The mating behaviour of *Saccharomyces* yeasts is controlled by the mating type locus (*MAT*) which has two alleles (*MAT a* and *MAT α*). Cells expressing *MAT a* (mating type a) conjugate readily with cells expressing *MAT α* (mating type α) to create *MAT a /MAT α* diploids. In heterothallic (*ho*) strains the mating types are stable and interconvert only rarely (10^{-6}) while in homothallic (*HO*) strains interconversion of mating types occurs as often as every cell division³. A haploid *HO* spore of one mating type can divide twice to yield four daughter cells, two exhibiting the parental mating type and two of the opposite cell type. Progeny of opposite mating type then conjugate to establish *MAT a /MAT α* diploid cell lines in which both switching and mating functions are shut off. This switching process is due to heritable but reversible alterations of the *MAT* locus, and the continued presence of the *HO* gene is not required for the maintenance of the altered state of that locus⁴⁻⁷.

Two other genetic loci, *HMA* and *HMA α* are also required for switching⁸⁻¹⁰. The genetic nomenclature originally associated with these loci has become somewhat confusing in the light of the function provided by these genes, as described below. We will use here a simpler system which was proposed at the International Congress on Yeast Genetics in 1978. Both forms are presented in Table 1. The *HML* locus (formerly *HMA*) on the left arm of chromosome III as it is conventionally drawn (see Fig. 1), has two naturally-occurring co-dominant alleles: *HML a* and *HML α* . Similarly, the *HMR* locus on the right arm of the chromosome has two codominant alleles *HMR a* and *HMR α* . Either *HML a* or *HMR a* is required for the activation of the a mating type gene at the mating type locus¹¹. Similarly, either *HML α* or *HMR α* is required for the activation of the α mating type gene at the mating type locus¹².

As an explanation of the heritable changes observed at the *MAT* locus during switching, Hicks, Strathern and Herskowitz¹ proposed the 'cassette model' for mating type interconversion in which changes of cell type occur through transposition of a and α mating type genes to the *MAT* locus from unexpressed storage sites at *HML* and *HMR*. This model provides a molecular explanation of the heritability of the switched *MAT* alleles and accounts for the striking observation that cells carrying a mutant *mata⁻* (sterile) allele can switch to *MAT a ⁺* and then 'healed' to *MAT α ⁺* (fertile) within three generations in the presence of the *HO* gene. This healing occurs with a variety of *mata⁻* and *mata⁻* mutations¹³⁻¹⁵. Therefore, cells capable of switching must contain reserve copies of *MAT a* and *MAT α* information in the genome. In the cassette model, the *HML* and *HMR* loci are proposed to contain normally unexpressed copies of the mating type gene; the *HML α* and *HMR α* alleles corresponding to α information and the *HML a* and *HMR a* alleles to a information. A change of cell type is thus viewed as a loss of previously expressed DNA from *MAT* concomitant with an insertion of a replica (cassette) of a or α information derived from the *HML* or *HMR* loci. This scheme is shown in Fig. 1. Several predictions of the cassette model have been recently verified: (1) a number of mutations have been characterised which apparently lead to the expression of a or α information at the normally silent *HML* and *HMR* loci¹⁵⁻¹⁸; (2) chromosome rearrangements have been identified which lead to the activation of *HML α* or *HMR α* by fusion to *MAT⁺*¹⁹; and (3) mutations in the *HML* and *HMR* loci have been transposed to *MAT* by switching^{20,21}.

Table 1 Nomenclature of genes involved in switching *MAT*^{*}

Old	New
<i>HMA</i>	<i>HMLα</i>
<i>hma</i>	<i>HMLa</i>
<i>HMAα</i>	<i>HMRa</i>
<i>hma</i>	<i>HMRα</i>

Capital letters are used for both alleles of each locus to indicate codominance. Recessive mutant alleles are indicated by lower case designations (for example, *hmra⁻*).

^{*} Data compiled from Harashima *et al.*⁸ and Naumov and Tolstorukov⁹.

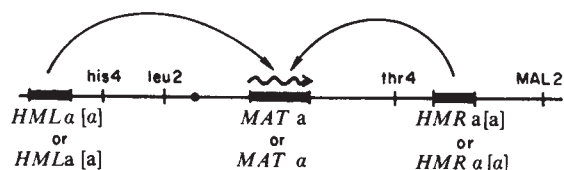


Fig. 1 Map of chromosome III and the cassette model for mating type interconversion. *HML* and *HMR* are sites which contain unexpressed copies of the *a* and α genes. Most laboratory strains exhibit the *HMLa* and *HMRa* arrangement of storage genes. During interconversion a copy of one of the silent genes is transposed to the *MAT* locus region where it is expressed. The fate of the original *MAT* allele is not clear but it has not been shown to reappear in the genome and is presumably lost.

Identification of a cloned α gene by transformation in yeast

The cassette model makes strong predictions concerning the physical arrangement of the *a* and α DNA sequences in the genome. Clearly, if this model is correct, an α gene at *MAT* should have an homologous counterpart at *HMLa* and *HMRa*. Furthermore, as *a* and α are distinct codominant genes, there must be some physical difference between the *MATa* and *MATa* alleles. The cassette model predicts that a similar difference should distinguish the *HMLa* and *HMLa* alleles and the *HMRa* and *HMRa* alleles. To test these predictions, it was necessary to identify and clone a mating type gene and use it as a hybridisation probe to examine the arrangement of homologous *a* and α sequences in the genome. As there is no easily obtainable physical probe for a yeast mating type gene, we used recently developed techniques for yeast transformation to isolate a mating type α gene by its ability to complement a defective mating type locus in yeast.

The vehicle that we have constructed for cloning yeast genes consists of a 1.4×10^6 molecular weight *EcoRI* fragment of the B form of *ScpI* cloned into the *EcoRI* site of the bacterial vector pBR322 along with a *PstI* fragment containing the yeast *LEU2* gene inserted into a *PstI* site in *ScpI* sequence. This plasmid, designated YEp13 in accord with the nomenclature of Davis *et al.*²², confers resistance to ampicillin and tetracycline in *Escherichia coli*, complements the bacterial *leuB6* and yeast *leu2* mutations, replicates as a plasmid in both organisms and transforms yeast at high frequency²³. A library of cloned yeast DNA molecules was constructed by ligating *BamHI* restriction fragments from α haploid strain DC56 into the unique *BamHI* site in the Tet^R determinant of YEp13 and transforming *E. coli* C600 and screening for Amp^R Leu⁺ Tet^S colonies. These colonies containing cloned yeast DNA were divided into subsets of 200 yeast clones each and the clones were amplified by standard procedures as described previously²⁴.

To detect a plasmid containing the yeast α mating type gene, pooled plasmid DNA was used to transform a yeast strain (DC49) containing a defective *mata*⁻ allele along with a non-reverting double mutation in the *leu2* gene²⁵. The recessive *mata*⁻ allele was used so that transformants containing an α plasmid would exhibit the α mating phenotype (*MATa*/*mata*⁻ = α). Transformation was performed as described previously²⁵ and Leu⁺ colonies were screened for α mating phenotype.

One subset of the clone library yielded Leu⁺ transformants of DC 56 with the α mating phenotype. Several such isolates were purified as single colonies and DNA was prepared from each by the method of Cryer *et al.*²⁶. The isolated DNA was used to re-transform the original bacterial strain (C600) to Amp^R Leu⁺, and plasmid DNA was prepared from the bacterial transformants. When the plasmid isolated from this bacterial strain was used to re-transform yeast strain DC49, 100% of the Leu⁺

transformants exhibited the α phenotype. This plasmid, designated YEp13-7, was shown by restriction mapping to consist of the original vector plus an insert of four *BamHI* restriction fragments totalling about 16 kilobase pairs. We demonstrated that only one of those *BamHI* fragments is required for the α phenotype by subcloning it on YEp13 and transforming DC49. This 6.6-kilobase pair yeast fragment is designated *Bam26* and is outlined in Fig. 2. The α Leu⁺ transformants of DC49 with YEp13-26 (YEp13 plus *Bam26*) were subsequently crossed with *a leu2*⁻ haploid tester strains. The resultant diploids were able to undergo meiosis and sporulation, indicating that the YEp13-26 plasmid supplies not only the functions necessary for α mating ability but the α sporulation functions as well. A restriction map of the insert in YEp13-26 is presented in Fig. 2. Leu⁻ mitotic segregants from these strains (the result of loss of the plasmid) concomitantly lost both α mating and sporulation functions.

Physical mapping of the *MAT* locus

The results described above were consistent with the notion that plasmid YEp13-26 contains a functional α gene but would have also been obtained if the cloned insert somehow disturbed the normal regulatory mechanisms in the transformed cell, allowing an unexpressed copy of α (that is, the copy proposed to be present at *HMLa*) in the DC49 genome to be expressed constitutively. Fortunately, it was possible to identify the cloned DNA as a nucleotide sequence homologous to the mating type locus by mapping the naturally-occurring differences in restriction patterns between *a* and α strains. This restriction fragment polymorphism was observed when the YEp13-26 plasmid was used as a hybridisation probe against Southern blots of restricted DNA from a series of different yeast strains. By crossing strains carrying different versions of a polymorphic restriction fragment and analysing DNA from individual meiotic segregants, the size difference (and hence the cloned DNA sequence used as a probe) can be mapped against known genetic markers. This strategy has previously been used to map a tRNA clone to the *SUP4* locus of yeast²⁷ and we have used it to identify a clone containing the yeast *can1* gene²³.

The α/a diploid strain XJ777 used for mapping the cloned α gene was made by crossing the haploid strain DC5 (*a*) and DC55 (α). The complete genotypes of the strains are presented in Table 2. Tetrads were dissected and genetic markers scored by standard techniques²⁸. DNA was then isolated from the parent strains and from individual spore clones derived from seven tetrads, digested with restriction endonuclease and displayed on agarose gels for Southern blotting²⁹. The blots were probed with ³²P-labelled DNA from plasmid pBR322-26.2 which contains the larger *EcoRI*-*BamHI* subfragment of YEp13-26 (Fig. 2). Figure 3 is an autoradiograph of a blot containing two representative tetrads from XJ777 digested with restriction endonuclease *HindIII*. There are two striking features of the hybridisation pattern. First, the hybridisation probe, which contains a *HindIII* site and therefore might be expected to hybridise to two chromosomal *HindIII* fragments, actually hybridises strongly to four such fragments. Second, two of the

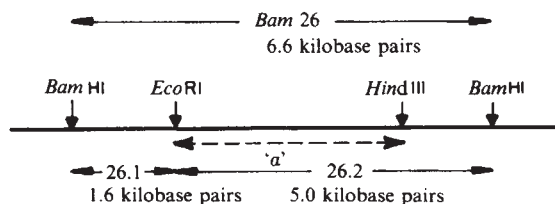


Fig. 2 Restriction map of the *BamHI* fragment containing α functions. The dotted arrow represents the limits of the α sequence which are homologous with restriction fragments containing *MAT* and *HMR*.

Table 2 Genotypes of yeast strains used

DC49	<i>HMLα leu2-3, 2-112 matα HMRα his3 can1</i>
DC5	<i>HMLα leu2-3, 2-112 MATα HMRα his3 can1</i>
DC55	<i>HMLα MATα HMRα lys2 arg4 tyr1</i>
DC89	<i>HMLα his4 leu2 matα thr4 hmrα met13 lys1 trp1</i>
	<i>HMRα + + matα + hmrα met13 lys1 trp1</i>
DC90	<i>hmlα his4 leu2 matα thr4 HMRα MAL2 lys1 met13 ade6</i>
	<i>hmlα + + matα + HMLα mal12 + + +</i>
X10-1B	<i>HMLα MATα HMRα HO his5 ade5 met4 met13 ura4</i>
	<i>HMLα MATα HMRα HO his5 ade5 met4 met13 ura4</i>
JH204	<i>HMLα MATα HMRα HO his5 ade5 met4 met13 ura4</i>
	<i>HMLα MATα HMRα HO his5 ade5 met4 met13 ura4</i>
JH210	<i>HMLα MATα HMRα HO his5 ade5 met4 met13 ura4</i>

four *Hind*III fragments exhibit a length polymorphism which segregates in mendelian fashion. The polymorphic band C exhibits 100% co-segregation with mating type. All seven tetrads had the parental arrangement of mating type relative to the band C polymorphism (complete data not shown). This band shift, linked to the *MAT* locus, which is consistent with a size difference of ~200 base pairs, was also observed when tetrad DNA was cut with *Eco*RI or *Bam*HI. In each case the fragment corresponding to the *MAT α* allele seemed to be ~200 base pairs smaller than the *MAT α* fragment. Thus, we conclude that the insert in YEp13-26 is homologous to a DNA restriction fragment tightly linked to the *MAT* locus, providing strong additional evidence that this plasmid contains an α structural gene. The segregation of the band A polymorphism is not correlated with mating type and reflects a naturally occurring strain difference at a locus unlinked to *MAT* as described below.

The observation that the *MAT α* -containing restriction fragment is ~200 base pairs larger than the fragment containing *MAT α* when a variety of restriction enzymes are used indicates that the α and α strains used in this mapping differ by a small net insertion at or near the *MAT* locus. We demonstrated that this net size difference reflects a difference in the size of the *MAT α* and *MAT α* genes by taking a heterothallic yeast strain through a series of changes of cell type: *MAT α* to *MAT α* to *MAT α* to *MAT α* . DNA was made from each step and analysed as in Fig. 3. The switches from α to α resulted in 200-base pair increases in the *Hind*III fragment containing *MAT* and the α to α switch resulted in a net 200-base pair decrease (data not shown).

Mapping of *HMR* and *HML* fragments

Two of the additional hybridisation bands in Fig. 3 may be explained in terms of the cassette model; that is, they should represent restriction fragments corresponding to the *HML* and *HMR* loci. We therefore used the same mapping procedure to determine the identity of these bands. As *HML* and *HMR* are normally unexpressed, mapping these loci requires the construction of special strains in which their genotypes can be scored. We have used the *mar1*⁻ allele, first identified by Klar *et al.*¹⁶, to construct mapping strains in which the *HML* and *HMR* alleles are expressed. The *MAR1* gene product has been proposed to keep the information at the *HML* and *HMR* silent by negative control. In the *mar1*⁻ background, the *HML* and *HMR* loci act as if they contain active mating type genes. A normal haploid strain of genotype *HML α MAT α HMR α* becomes a non-mater in the presence of *mar1*⁻ because both α and α mating type genes are expressed. A strain of genotype *mar1*⁻ *HML α mat α hmr α* , on the other hand, behaves as an α -mater because *HML α* is expressed and the other two loci are defective. In such a strain, the α mating type maps to the *HML* locus, distal to *his4* (see Fig. 1) rather than to *MAT* (unpublished results). By constructing *mar1*⁻ strains with the appropriate arrangement of mutant and normal mating type alleles, as shown in Table 1, it is possible to place the only functional mating type genes at *HML* (*HML α /HML α* , strain DC89) or at *HMR* (*HMR α /HMR α* , strain DC90). The segregation of the α and α phenotypes in these strains thus reflects the segregation of the alleles at *HML* (DC89) or *HMR* (DC90) rather than *MAT*. Segregants from

these strains can then be screened by Southern blot analysis as described above. The strategy used in the construction of these mapping strains will be described elsewhere (J.H., J.N.S. and A.J.S.K., in preparation).

In the cassette model *HML* and *HMR* are sources of DNA sequences that become substituted into *MAT*. If, as suggested, *HMR α* is an unexpressed α sequence and *HMR α* is unexpressed α , then a strain segregating those alleles (DC90) should exhibit a ~200-base pair polymorphism as demonstrated for *MAT*. Specifically, *HMR α* and *HML α* should be ~200 base pairs larger than their *HMR α* or *HML α* alleles. Figure 4 shows autoradiographs of two tetrads derived from diploids in which *HML* or *HMR* has been made heterozygous for active α and α regions. The genotypes of these diploids are presented in Table 2. Once again, alleles at these loci are associated with a restriction fragment polymorphism. In Fig. 4a the upper band exhibits the polymorphism co-segregating with the *HML α* and *HML α* alleles. As in the case of the *MAT* locus, the α (*HML α*) allele

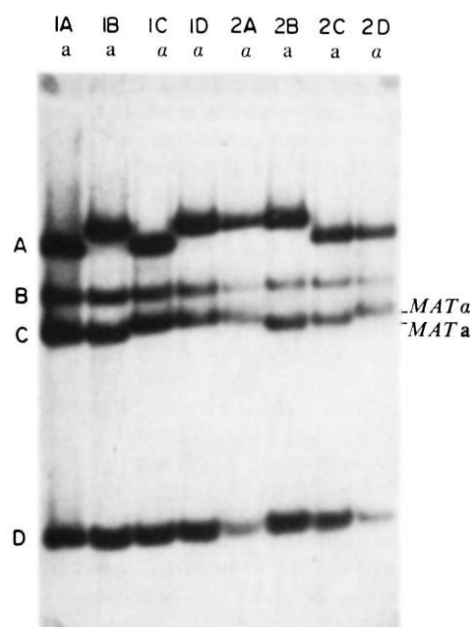


Fig. 3 Autoradiograph of two meiotic tetrads derived from crossing DC5 (α) and DC55 (α). DNA was cut with *Hind*III, blotted on nitrocellulose and hybridised with pBR322-26.2. The size shift in band C correlates with the segregation of mating type. Yeast culture media and standard yeast techniques have been described previously^{3,7}. Yeast strains used in this work are presented in Table 1. Plasmids were propagated in *E. coli* B strain C600 (*leuB thr pro*) grown in Luria broth supplemented with ampicillin (100 μ g ml⁻¹) or tetracycline (25 μ g ml⁻¹) as required or on M9TP (1XM9 salts³⁰, 1% glucose, 1 mM MgCl₂, 20 μ g ml⁻¹ thiamine, 100 μ g ml⁻¹ threonine, and 10 μ g ml⁻¹ proline). *E. coli* strains were grown in M9CAA (1XM9 salts, 1% casaminoacids, 1% glucose, 1 mM MgCl₂, and 20 μ g ml⁻¹ thiamine) for large scale plasmid preparations. Yeast DNA was isolated from appropriate strains by the method of Cryer *et al.*²⁶ using equilibrium centrifugation in CsCl as a final purification step. Plasmid DNA was isolated from *E. coli* by standard procedures. Restriction endonuclease digestions, DNA ligation reactions and DNA labelling with ³²P were performed as described previously²³. Fractionation of DNA was accomplished by electrophoresis on horizontal agarose slab gels poured and run in 40 mM Tris-acetate, pH 7.6, 1 mM EDTA, and 1.0 μ g ml⁻¹ ethidium bromide. Transfer of DNA from agarose gels to nitrocellulose filters was affected by the method of Southern²⁹. Hybridisation of labelled DNA to DNA immobilised on nitrocellulose filters and transformation of yeast cells with purified plasmid DNA was performed as described elsewhere^{24,25}. Transformation of bacterial cells with DNA was performed using calcium shocked *E. coli* cells as described elsewhere³¹.

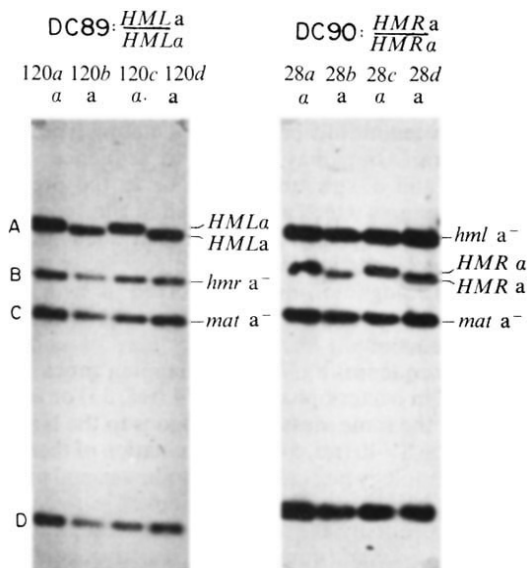


Fig. 4 Autoradiographs of tetrads from diploids DC89 and DC90 showing segregation of size shifts in bands A and B correlated with heterozygosity for *HMLa/HMLa* and *HMRa/HMRa*, respectively. DNA was digested with *Hind*III and pBR322-26.2 was used as a probe. Approximate sizes of the *a* and *α* bands in each pair (in kilobase pairs) are: A, 5.8 and 6.0; B, 4.6 and 4.8; C, 4.1 and 4.3.

appears to be ~200 base pairs smaller than the corresponding *α* (*HMLα*) allele. This size variation segregated in the parental configuration relative to mating type in 12/12 meiotic clones. Similarly, when the *HMR* locus is heterozygous, the size variation can be seen in the second largest band (Fig. 4b). Once again, the ~200-base pair larger version co-segregates with the *α* phenotype, that is, the *HMRα* allele (8/8 meiotic clones tested). The smallest restriction fragment (band D) corresponds to the small partial *Hind*III fragment carried on the right-hand end of the *Bam*26 fragment (see Fig. 2) and remains uniform in size.

On the basis of these data, we conclude that the cloned *α* mating type gene is at least partly homologous with all three loci (*MAT*, *HML*, *HMR*) proposed to contain mating type genes. This homology is observed irrespective of the allele (*a* or *α*) present at those loci. Furthermore, because a similar size variation is seen at each locus and because such a size difference appears in digests from several different restriction enzymes, we believe that this polymorphism represents an intrinsic size difference between the *a* and *α* alleles.

Location of the cloned *α* gene

Because the cloned fragment YEp13-7 hybridises to *MAT*, *HML* and *HMR*, it is not possible from the above observations to determine whether the *α* gene comes from *MATα* or *HMLα*. Although *HMLα* would not be expected to be expressed in a *MAR*⁺ strain such as DC49, it is possible that, as an artefact of the cloning procedure, *HMLα* could have been released from the normal mechanism of negative control. In fact, several lines of evidence indicate the *α* gene identified by function in YEp13-26 is the one originating from *HMLα*: (1) *Bam*HI fragment 26 (Fig. 2) hybridises more strongly to the restriction fragment identified as *HML* than to *MAT* or *HMR*; (2) the *α* phenotype of *mar1*⁻ strains transformed with YEp13-26 is more pronounced than in *MAR*⁺ strains; and (3) the *Bam*HI-*Eco*RI subclone 26.1 (Fig. 2) does not hybridise to the *Hind*III fragments from total yeast DNA identified as containing *MAT* or *HMR* but does hybridise to the *Hind*III fragment identified as containing *HML* (data not shown).

Transposition in homothallic switching

From the observations described above, we predicted that the efficient interconversion of mating types observed in homothallic yeast would also involve insertion and excision of DNA at *MAT*. To test this prediction we monitored the effect of a conversion from *a* to *α* in a homothallic strain. A homothallic (*HO*) haploid normally forms a stable *a/α* diploid cell line through the process of mating type interconversion and mating among the progeny. Once in the *a/α* state, the *HO* switching functions no longer operate and the mating type alleles are stable. However, it is also possible to obtain stable haploid strains containing a functional *HO* gene through manipulation of the states of the *HML* and *HMR* loci. An *HO HMLα MATα HMRα* strain, for example, does not switch and remains an *α* haploid. The cassette model proposes that such a strain appears stable because it lacks a copy of *a* information. It is therefore possible to construct a strain which has two possible developmental fates. If a cell of genotype *HO HMLα MATα HMRα* is allowed to divide repeatedly, two cell types are obtained: cells with the original *a* phenotype whose genotype is unchanged, and *α* cells which now have the *HO HMLα MATα HMRα* genotype. These *α* cells are no longer capable of switching because they lack a copy of the *a* gene, and, if experimentally isolated from the sibling *a* cells, will grow into a stable *α* haploid clones. The *a* cells, of course, can still switch, giving rise to addition *a* and *α* progeny which can mate among themselves to form stable *a/α* diploid cell lines. According to our interpretation of the hybridisation patterns described above, the switching events responsible for the creation of the two distinct cell lines described above should be reflected in the sizes of the restriction fragments carrying mating type genes. The *HO HMLα MATα HMRα* haploid should have only a single *MATα* band, while the diploid *HMLα/HMLα MATα/MATα HMRα/HMRα* should contain two *MAT* bands, one derived from the parental *a* allele and one from the newly inserted *α* allele. Thus we should be able to correlate a change in cell type with a change in the structure of the chromosome in the progeny of a single cell.

The results of such an experiment are shown in Fig. 6. The strains used were created by repeated back crossing (3X) of strains carrying particular arrangements of the *HML* and *HMR* alleles with the normal homothallic diploid X10-1B. X10-1B and JH210 (Fig. 6 lanes *a*, *b*) are stock strains, while lanes *c* and *d* contain DNA from sibling progeny of a single *HO HMLα MATα HMRα* spore from strain JH204 (Fig. 5). The isolation of these strains followed the strategy described above and used techniques for the pedigree analysis of individual yeast cells described previously⁷. As shown in Fig. 6, the size variation in

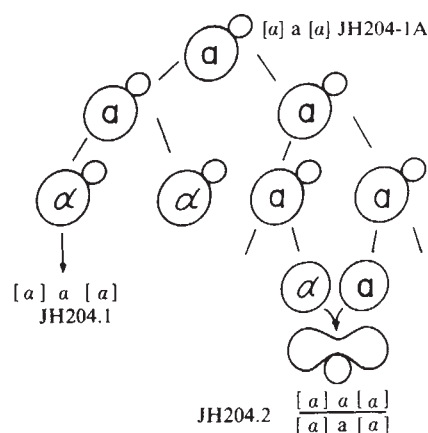


Fig. 5 Schematic of pedigree analysis used to isolate homothallic strains of different developmental potential from a single haploid spore.

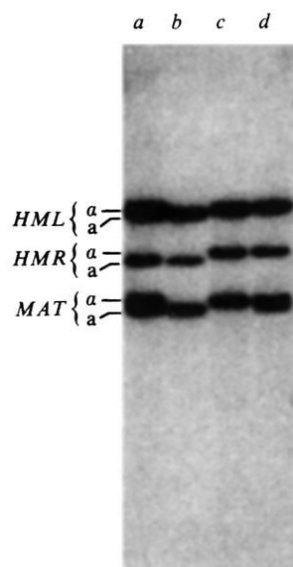


Fig. 6 Autoradiograph of *Hind*III digested DNA from related homothallic strains probed with a subfragment of *Bam*26 consisting of the 3.6-kilobase pair *Eco*RI-*Hind*III region (see Fig. 1). Lane *a* is the *a*/*α* diploid X10-1B. Lane *b* is an *HMRα* haploid strain, JH210. Lanes *c* and *d* represent haploid and diploid derivatives (JH204.1 and JH204.2) of a single haploid spore from JH204.

the *HML*, *HMR* and *MAT* bands follows the genotype of each strain as determined by genetic criteria. Furthermore, the sibling strains (lanes *c* and *d*) which have undergone a homothallic switching event and, thus, a change in developmental potential, show a difference in the banding pattern reflecting the condition of the *MAT* locus. In the haploid *α* strain (lane *c*) a single *MATα* band is present, while in the heterozygous diploid (lane *d*) a doublet consisting of one *MATa* and one *MATα* size band appears. It is impossible in such an experiment to trap the original *HO HMLα MATa HMRα* cell as a haploid as it continually switches to *α* and mates. However, we can infer the size of the original *MATa* band in that spore from the *MATa* bands present in the closely related strains X10-1B and JH210 (lanes *a* and *b*).

Discussion

We have identified and cloned a functional *α* mating type gene by transformation in yeast. The cloned sequence was identified on the basis of its ability to complement a defect at the *MAT* locus for both mating and sporulation functions. Through subsequent hybridisation analysis we have shown that sequences homologous to the cloned gene exist at three locations in the yeast genome at sites near the *HML*, *HMR* and *MAT* loci. These observations are consistent with the cassette model for mating type interconversion in which the *HML* and *HMR* loci are proposed to contain unexpressed storage copies of the *α* and *a* mating type genes. The sizes of the three restriction fragments corresponding to these loci depend on the mating type allele present at each locus as defined by genetic criteria. Fragments containing the *α* allele uniformly appear to be ~200 base pairs larger than corresponding fragments containing *a*. Furthermore, using this apparent size difference, we have shown that both homothallic and heterothallic interconversion of mating types involves a change in size of the restriction fragment mapped to *MAT*. This change in restriction pattern provides a physical demonstration of DNA rearrangement during the interconversion process. We interpret this alteration in restriction pattern as being a consequence of transposition of mating type genes from *HML* and *HMR* to the *MAT* locus.

A fundamental prediction of the cassette model is that the cloned *α* gene should exhibit strong homology to all copies of *α* proposed to exist in the genome (*HMLα*, *MATα*, *HMRα*). However, we have observed that this cloned sequence hybridises strongly to genetic loci harbouring *a* mating type alleles as well. This hybridisation may be due to sequence similarity between the *a* and *α* structural genes or to the presence of recognition sequences which are involved in site specific transposition. We had originally proposed that such recognition sequences might be responsible for the specificity and directionality of the recognition events involved in switching^{1,2,32}. However, it is also possible that, although the *a* and *α* gene products are functionally distinct, they may be coded from closely-related sequences either as overlapping genes similar to those described in bacteriophage ϕ X174 (ref. 33) or as processed products of the same message analogous to the large T and small t proteins in SV40 (ref. 34). Determination of the source of the observed homology must await heteroduplex and nucleotide sequence data on the cloned *a* and *α* genes.

The results presented here confirm the physical predictions of the cassette model and, coupled with existing genetic data, provide strong evidence for transposition of structural genes as a developmental regulatory mechanism. However, the molecular details of the mechanism remain unclear. As the genetic content of the *HML* and *HMR* loci is not altered during switching, it is clear that the transposition event is not reciprocal. Therefore switching must involve both replication of the source gene and substitution of a copy of that gene for the one originally present at *MAT*. It is possible that copies of the source loci are replicated and released as independent molecules which then diffuse to the *MAT* locus and undergo a separate recombination event resulting in substitution. We have looked for a 'diffusible cassette' using Southern blots of undigested DNA from strains constitutive for switching. Such molecules have not been observed in conditions where one molecule per cell would have been easily detectable. Alternatively, replication and substitution need not be separate events but could be the result of a concerted process involving normal recombinatorial intermediates. As we have described elsewhere³², mating type interconversion is functionally equivalent to unidirectional gene conversion. As substantial homology exists between the source loci and *MAT*, conversion could occur through the same steps proposed for gene conversion at other genetic loci with the proviso that during heteroduplex repair, the source sequences are always maintained. Such unidirectional gene conversion has been observed for certain alleles in both *Sordaria brevicolis*³⁵ and *Schizosaccharomyces pombe*³⁶.

The notion that switching might be the result of a modification of normal recombination mechanisms is strengthened by the observation that the *HML*, *HMR* and *MAT* loci are unusually susceptible to intrachromosomal recombination events leading to deletions³⁷ or chromosomal rearrangements¹⁹. Genetic selections for rate heterothallic *α* to *a* interconversions often yield a specific deletion extending from *MATα* to *HMRα* (see Fig. 1). This deletion apparently removes the *MATα* gene and leads to the constitutive expression of *HMRα*. The converse fusion of *MATa* to *HMLα* yields a circular chromosome containing the centromere in which *MATa* has been removed and *HMLα* is expressed¹⁹. These events may be due to pairing of homologous sequences followed by generalised reciprocal recombination or may represent abortive resolution of intermediates generated during site-specific transposition.

Although our results show that switching from *a* to *α* involves the net addition of a ~200-base pair segment to the *MAT* restriction fragment, the actual size of the transposed region has not been determined. One possible interpretation of this result is that the expression of the *α* and *a* mating type genes at *MAT*, *HML* or *HMR* depends solely on the presence or absence, respectively, of a 200-base pair translocatable element. Recent genetic results, however, make this possibility unlikely. Klar and Fogel²¹ and Kushner *et al.*²⁰ have shown that mutant information at the *HML* and *HMR* loci can be transposed to *MAT*.

during both homothallic and heterothallic switching. Therefore some sequences present in the α and α structural genes must be transposed during switching. Our results further indicate that the α transposon is less than 3.6 kilobase pairs as described in Fig. 2 legend.

We have demonstrated that interconversion of mating types in *Saccharomyces* involves an alteration of the DNA sequence at the mating type locus of this organism. We believe that this genetic change is the result of sequential transposition of α and α structural genes to *MAT* from the *HML* and *HMR* storage loci. This change is heritable but is also subject to subsequent transposition events. These interconversions of cell type occur in an orderly fashion: (1) they occur in a defined subset of clonally-related cells^{3,7}; and (2) cells do not randomly become α or α ;

they are directed to change to the new cell type by functions under the control of mating type alleles^{3,7,14}. Thus, this simple mating type switch exhibits many of the properties observed in the differentiation of cell types in more complex organisms and the mechanism of mating type interconversion provides a paradigm for the study of other developmental events.

We thank Carolyn McGill, Jean McIndoo and Sajida Ismael for technical assistance, James Broach for sharing his technical expertise and suggestions, John Roth for helpful criticism and Marie Moschitta for preparation of the manuscript. This work was supported by NIH grants GM 25678 (A.K.) and GM25624 (J.B.H. and J.N.S.) and NSF grant PCM78-07793 (J.B.H. and J.N.S.). Experiments involving molecular cloning were carried out under the provisions of the NIH guidelines.

Received 6 August; accepted 19 September 1979.

- Hicks, J., Strathern, J. & Herskowitz, I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A., Shapiro, J. & Adhya, S.) 457-462 (Cold Spring Harbor, New York, 1977).
- Herskowitz, I., Strathern, J., Hicks, J. & Rine, J. in *Proceedings of the 1977 UCN-UCLA Symposium: Molecular Approaches to Eukaryotic Genetic Systems* (eds Wilcox, C., Abelson, J. & Fox, C. F.) 193-202 (Academic, New York, 1977).
- Strathern, J. & Herskowitz, I. *Cell* **17**, 371-381 (1979).
- Winge, O. & Roberts, C. *C. R. Lab. Carlsberg, Ser. Physiol.* **24**, 341-345 (1949).
- Hawthorne, D. C. *Abstr. Proc. 11th Int. Congr. Genet.* **1**, 34-35 (1963).
- Oshima, Y. & Takano, I. *Genetics* **67**, 327-335 (1971).
- Hicks, J. & Herskowitz, I. *Genetics* **83**, 245-258 (1976).
- Harashima, S., Nogi, Y. & Oshima, Y. *Genetics* **77**, 639-650 (1974).
- Naumov, G. I. & Tolstoukov, I. I. *Genetika* **9**, 82-91 (1973).
- Harishima, S. & Oshima, Y. *Genetics* **84**, 437-451 (1976).
- Klar, A. & Fogel, S. *Genetics* **88**, 407-416 (1977).
- Arima, K. & Takano, I. *Genetics* (in the press).
- Hicks, J. & Herskowitz, I. *Genetics* **88**, 373-393 (1977).
- Strathern, J., Blair, L. & Herskowitz, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3425-3429 (1979).
- Klar, A., Fogel, S. & Radin, D. *Genetics* (in the press).
- Klar, A., Fogel, S. & MacLeod, K. *Genetics* (in the press).
- Rine, J., Strathern, J., Hicks, J. & Herskowitz, I. *Genetics* (in the press).

- Haber, J. E. & George, J. P. *Genetics* (in the press).
- Strathern, J., Newlon, C., Herskowitz, I. & Hicks, J. *Cell* **18**, 309-319 (1979).
- Kushner, P., Blair, L. & Herskowitz, I. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5264-5268 (1979).
- Klar, A. & Fogel, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4539-4543 (1979).
- Struhl, K., Stinchcomb, D. J., Scherer, S. & Davis, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1035-1039 (1979).
- Broach, J., Strathern, J. & Hicks, J. *Gene* (in the press).
- Hicks, J. & Fink, G. *Nature* **267**, 263-265 (1977).
- Hinnen, A. H., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 170-174 (1978).
- Cryer, D. F., Eccleshall, R. & Marmur, J. *Meth. Cell Biol.* **12**, 39-44 (1975).
- Goodman, H., Olsen, M. & Hall, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1977).
- Mortimer, R. K. & Hawthorne, D. C. in *The Yeasts* Vol. 1 (eds Rose, A. H. & Harrison, J.) 385-460 (Academic, New York, 1969).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Miller, H. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1974).
- Petes, T. D., Broach, J. R., Wensink, P. C., Hereford, L. M., Fink, G. R. & Botstein, D. *Gene* **4**, 37-49 (1978).
- Hicks, J. & Strathern, J. *Brookhaven Symp.* **29**, 233-242 (1977).
- Barrell, B. G., Air, G. M. & Hutchinson, C. A. III *Nature* **264**, 34-41 (1976).
- Pancha, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **15**, 2165-2169 (1978).
- McDonald, M. V. & Whitehouse, H. L. K. *Genetical Research* (in the press).
- Gutz, H. *Genetics* **69**, 317-337 (1971).
- Hawthorne, D. C. *Genetics* **48**, 1727-1729 (1963).

letters

A radio jet in SS433

R. E. Spencer

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire, UK

The peculiar emission line object SS433 (refs 1-3) has variable radio emission and an angular size of <0.1 arc s at centimetric wavelengths^{4,5}. The observations reported here show that the object at long radio wavelengths consists of this compact core and also a jet ~ 1 arc s long aligned in the same position angle as the extended structure in the associated supernova remnant W50. Both components vary in flux density.

The observations were made at 408 MHz with radio-linked interferometers consisting of the MK IA 76 m or the MK II 25 m $\times$ 37 m telescopes at Jodrell Bank and the 25 m telescope at Defford. All telescopes were equipped with 1-h circularly polarised feeds. The baseline length was 127 km, giving a resolution limit of 0.2 arc s. The dates of the observations are shown in Table 1. The source was tracked for long periods at all except the first two epochs, enabling models of the radio structure to be fitted to the fringe amplitudes obtained. Short observations only were available on the first two dates, but these were

sufficient to allow the flux densities of an unresolved core and an extended component to be found.

A sample visibility curve (that of 10 August 1979) is shown in Fig. 1. Each point represents the average of four integrations of 30 s duration. Also shown is the predicted visibility from a model consisting of an unresolved core (<0.2 arc s diameter) with a flux density of 0.8 Jy plus a gaussian component of 0.9 Jy with dimensions of 1.0 arc s FWHM by <0.2 arc s and elongated in position angle 100° . Similar models were fitted on the other dates giving flux densities, lengths of the elongated component and position angles as shown in Table 1. The flux densities of the two components show significant variations while changes in size and position angle are within measurement errors. The range of resolution available was not sufficient to establish whether the jet is symmetric about the core or displaced by up to ~ 0.5 arc s in position angle 100° . In addition, lack of phase information gives rise to an ambiguity of 180° in position angle. Nevertheless, the effect of an elongated component is quite clearly seen in the data.

This jet-like feature is very similar to structures found in the central regions of some extragalactic radio sources such as 3C236 (ref. 6) and NGC6251 (ref. 7). It has been proposed that these regions are the sites of radio emission from beams of relativistic matter flowing out from a central nucleus. Such beams may therefore be present in SS433 though at a much lower power level and smaller linear scale. Indeed, beams have

Table 1 Details of observations on SS433

Instrument	Date	Total flux density (Jy)	Core flux density (Jy)	Extended component flux density (Jy)	Size (arc s)	Position angle
MK IA-Defford	27 May 1976	1.9 ± 0.3	1.3 ± 0.3	0.6 ± 0.4	—	—
MK IA-Defford	4 February 1979	1.5 ± 0.2	1.3 ± 0.2	0.2 ± 0.2	—	—
MK IA-Defford	6 June 1979	2.6 ± 0.2	2.2 ± 0.2	0.4 ± 0.2	1.5	99°
MK II-Defford	31 July 1979	1.1 ± 0.1	0.5 ± 0.1	0.6 ± 0.1	1.0	102°
MK II-Defford	9 August 1979	1.7 ± 0.1	0.9 ± 0.1	0.8 ± 0.1	1.2	98°
MK II-Defford	10 August 1979	1.7 ± 0.1	0.8 ± 0.1	0.9 ± 0.1	1.0	100°
MK II-Defford	11 August 1979	1.7 ± 0.1	0.8 ± 0.1	0.9 ± 0.1	0.9	97°
MK II-Defford	12 August 1979	1.8 ± 0.1	0.8 ± 0.1	1.0 ± 0.1	0.8	97°

been proposed^{8,9} to explain the moving optical emission lines in SS433 and Begelman *et al.*¹⁰ have further suggested that these beams supply energy to the supernova remnant W50. On single-dish maps^{11,12} W50 has the appearance of a filled spherical shell of emission $\sim 1^\circ$ in diameter with regions extending to $\sim 2^\circ$ in overall size in position angle $100 \pm 10^\circ$. The orientation of the jet found here (99°) suggests that these outer regions have been blown out by beams from SS433. The range of angular scales involved is $\sim 7,000:1$ and the beams must have maintained their orientation within 10° for at least the light travel time from SS433 to the outer regions of W50 which is ~ 200 yr if SS433 is the suggested distance of 3.5 kpc (ref. 3).

The moving optical emission shows a periodicity in velocity of 164 days and it has been proposed that beams responsible for the emission describe a cone of half-opening angle 17° inclined at 78° to the line of sight⁹. Such motion may be caused by precession of the parent object¹⁰. The phase of the precession is such that the position angle would be expected to change by 30° from 6 June to 12 August 1979, whereas the measurements show that any precession of the radio jet must be $\leq 5^\circ$ over this period. The length of the radio jet if SS433 is at a distance of 3.5 kpc is 5×10^{16} cm, about 100 times the size suggested for the moving regions of optical emission¹⁰ and so it is possible that the effects of precession on position angle are smoothed out within the radio jet.

The flux densities of both the core and jet varied, particularly in the interval between 6 June 1979 and 31 July 1979. The electron lifetime at 1 GHz, if equipartition between relativistic electron energy density and magnetic energy density is assumed in the jet, is $\sim 1,000$ yr and so the variability is more likely to arise from adiabatic expansion of small emitting regions in the jet than from radiative energy losses. The core variability may be associated with a changing angle between the beams and the line of sight as recently discussed for much more powerful radio sources¹³. Further study of the flux density variations of the core and jet over several periods and high-resolution VLBI maps will show if this is the case, and give information on relationships between the optical properties and the radio structure.

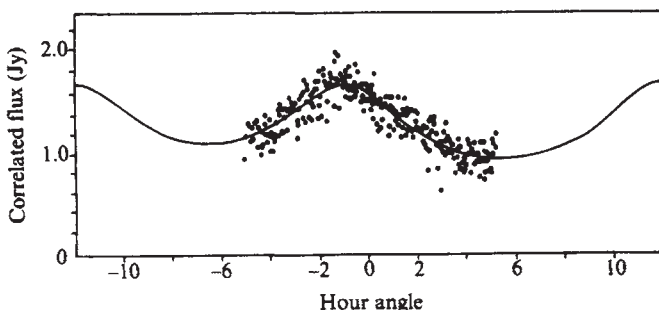


Fig. 1 Fringe amplitude versus interferometer hour angle on 10 August 1979. The smooth curve shows the expected fringe amplitude from a model consisting of a point source coincident with an extended gaussian component 1.0 arc s long to half maximum brightness and <0.2 arc s wide, with flux densities as given in Table 1.

I thank Dr Dennis Walsh for data and Owen Stewart for assistance.

Received 19 September; accepted 18 October 1979.

1. Stephenson, C. B. & Sanduleak, N. *Astrophys. J. Suppl.* **33**, 459–469 (1977).
2. Clark, D. H. & Murdin, P. *Nature* **276**, 44–45 (1978).
3. Margon, B. *et al. Astrophys. J. Lett.* **230**, L41–L45 (1979).
4. Seaquist, E. R., Gregory, P. C. & Crane, P. C. *IAU Circ. No.* 3256 (1978).
5. Ryle, M., Caswell, J. L., Hine, G. & Shakeshaft, J. *Nature* **276**, 571–573 (1978).
6. Fomalont, E. B. & Miley, G. K. *Nature* **257**, 99–102 (1975).
7. Readhead, A. C. S., Cohen, M. H. & Blandford, R. D. *Nature* **272**, 131–134 (1978).
8. Fabian, A. C. & Rees, M. J. *Mon. Not. R. astr. Soc.* **187**, 13P–16P (1979).
9. Abell, G. O. & Margon, B. *Nature* **279**, 701–703 (1979).
10. Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* (in the press).
11. Velusamy, T. & Kundu, M. R. *Astr. Astrophys.* **32**, 375 (1974).
12. Geldzahler, B. J., Pauls, T. & Salter, C. J. *Astr. Astrophys.* (in the press).
13. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34–48 (1979).

Low energy γ -ray spectrum of NGC4151

F. Perotti, A. Della Ventura, G. Sechi & G. Villa

Laboratorio di Fisica Cosmica e Tecnologie Relative, CNR Via Bassini 15, Milan, Italy

G. Di Cocco

Laboratorio TESRE, CNR Via De' Castagnoli 1, Bologna, Italy

R. E. Baker, R. C. Butler, A. J. Dean, S. J. Martin* & D. Ramsden

University of Southampton, Department of Physics, Southampton, UK

Following the first detection of X-ray emission from the Seyfert galaxy NGC4151¹, later observations^{2–4} have demonstrated that the X-ray spectrum up to 100 keV may be described by a power law having a differential spectral index $1 < \alpha < 1.6$ with a turnover at low energies caused by photoelectric absorption. Recently the results of an observation of this galaxy in the energy range 25–190 keV (ref. 5) have shown it to have a very hard spectrum with $\alpha \approx 0.9$. Here more detailed results of the original observation of NGC4151 (ref. 6) using the MISO telescope are presented.

The Seyfert galaxy NGC4151 was studied during the first balloon flight of the MISO telescope⁷ on the 22 May 1977. The instrument consisted of a Compton coincidence detector having in this flight an aperture of $3^\circ \times 20^\circ$ FWHM which was the same for both the 'single' and 'coincidence' γ -ray events. The telescope (energy range 0.04–20 MeV) was pointed using an alt-azimuth orientation system which had a precision of about 0.3° .

NGC4151 was observed at two different zenith angles (18° and 7°) over a period of ~ 3 h. Half of this time was devoted to

*Present address: Science Research Council, Appleton Laboratory, Slough, UK.

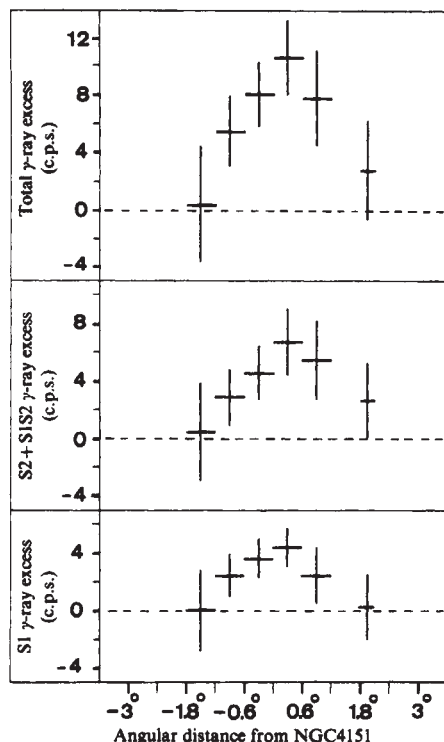


Fig. 1 γ -Ray excesses count rate (S1, S2+S1S2, total) versus the angle between the azimuthal axis of the MISO telescope collimator and the direction of NGC4151. The background measurement was made in a region of the sky some way ($>10^\circ$) from NGC4151.

background measurements. Observations at each zenith angle were made using an off-on-off source sequence. The source was allowed to transit through the field of view of the telescope and care was taken to ensure that no known X- and γ -ray sources were in the field of view during the background measurements. During the two observations of NGC4151 the float altitude varied between 6.2 and 8.0 mb but stayed close to 7 mb during the ON source observation. The geographical position remained practically constant during the entire observation period. To eliminate any systematic errors due to zenith variations of the background, the data were analysed separately for each zenith angle and the results statistically combined. No dependence of the counting rate on the azimuth angle was found. This means that, if an azimuthal effect exists, it is much less than the statistical errors, as expected from the fact that the entire observation was carried out at small zenith angles. Study of the housekeeping data and in flight γ -ray calibrations showed that there were no measurable changes in the gain of the system during the period. The only measurable systematic change of the background counting rate was due to altitude fluctuations. This variation, measured during the OFF source period, was found to be well described by a linear relationship with pressure for each detector element at both zenith angles. During the ON source period the background counting rate, computed from a knowledge of the balloon altitude, was subtracted from the counting rate generated by the source plus background for each channel. In this way the contribution from the source was obtained, giving a γ -ray excess of 3.0 ± 0.7 counts per second (c.p.s.) for the S1 detector (liquid scintillator), 3.1 ± 0.9 c.p.s. for the S2 detector (sodium iodide) and 1.1 ± 0.4 (c.p.s.) for the S1S2 Compton-coincidence system. The total γ -ray excess was 7.2 ± 1.2 c.p.s., which corresponded to a 6σ excess. This was considerably larger than the upper limit of the sum of all possible systematic errors. The correlation between the excess count rate and the angle between the azimuthal axis of the collimator and the direction of NGC4151 is shown in Fig. 1. The absence of data between 1.2°

and 1.8° in Fig. 1 is associated with a loss of recorded data. The lack of data just outside the 3° field of view of the MISO telescope is a consequence of the background measurements being made in a region of the sky some way ($>10^\circ$) from NGC4151. A χ^2 best fit of the angular profile data gives the γ -ray excess emission centred about a position $182.2^\circ \pm 0.8^\circ$ in RA, which effectively coincides with the position of the Seyfert galaxy NGC4151 (RA = 182.06°). The positional uncertainty associated with the second (zenith) axis is about $\pm 10^\circ$. The only other known X-ray source which came within the field of view during this observation was the BL Lac object 2A1219+305. However, as this object was at the edge of the aperture of the telescope (10% relative efficiency for photon detection) the proposed identification of the observed γ -ray excess with NGC4151 seems justified.

A technique similar to that used by Dolan⁸ was used to convert the energy loss spectrum in each of the three individual γ -ray channels S1, S2, S1S2 to a photon spectrum at the top of atmosphere. The absorption of photons in the residual atmosphere and the redistribution of photon energies through Compton interactions and pair production as well as the energy resolution of the telescope were taken into account. The precision of the system was estimated to be better than 10%. A statistical combination of all three individual spectra is shown in Fig. 2 together with some relevant X-ray measurements. For clarity, the measurements of OSO7 (ref. 2) and the 1976 Ariel 5

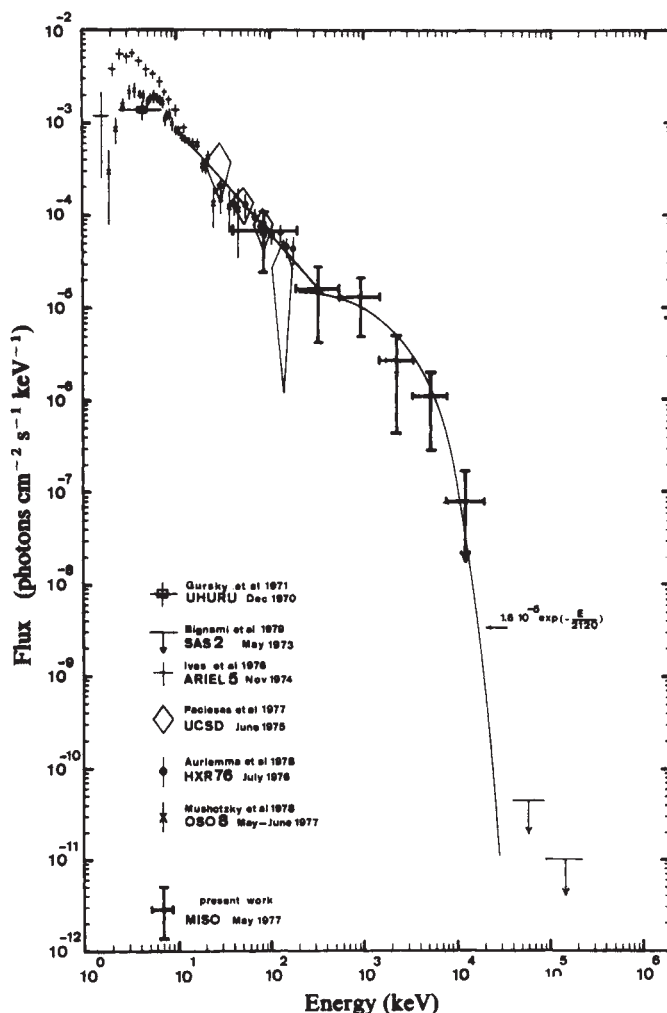


Fig. 2 The photon spectrum of NGC4151 obtained by the MISO telescope is shown together with some previous hard X-ray data. Also shown are the 2σ upper limits obtained by SAS2 at high energy¹¹.

data⁹ have not been shown as they are essentially identical to the more recent measurements of OSO8 (ref. 10). The figure does not show the low energy γ -ray results reported earlier by the Max Planck group¹¹ because of the uncertainty cast on the validity of this measurement by a more refined analysis of the data¹². Also shown are the 2σ upper limits set by SAS2 at higher energies¹³.

The Ariel 5 (ref. 14) and the OSO8 data provide evidence for the variability of NGC4151 on a time scale of about one week with changes in the flux of about a factor of 2. The observation with the MISO telescope was made on the 23 May 1977, a few days before the OSO8 observations. Inspection of the light curve from OSO8 suggests that the source was in 'high' state of activity during our observation period. The recently reported upper limits at MeV energies by the Rice group¹⁵ and by the University of California¹⁶ indicate that time variability in the γ -ray emission of NGC4151 must exist over a period of several months.

An attempt to fit our data with a single power law leads to a reduced $\chi^2 \approx 1.9$ with 4 d.f. The variability of this source is well established and the use of data taken at different epochs is not entirely satisfactory, however, when X-ray data above 20 keV (refs 4, 5, 10) are included, a similar treatment over the entire energy range leads to a reduced $\chi^2 \approx 2.3$ with 19 d.f. On the other hand, the data can be fitted better with a power law ($\alpha \approx 0.9$) up to 2.5–4 MeV followed by a break and a steeper spectrum ($\alpha = 3.4$ –4) at higher energies as suggested by both our data and the upper limits set by SAS2. The resulting reduced χ^2 value for this kind of fit is 0.73 with 19 d.f. Alternatively, our measurements may be represented at energies higher than 200 keV by an exponential of the form

$$dI/dE = (1.6 \pm 0.7) 10^{-5} \exp(-E/(2120 \pm 480))$$

photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$

corresponding to a reduced $\chi^2 \approx 0.5$ with 3 d.f. The data between 10 and 200 keV would then be best fit by a power law of the form

$$dI/dE = (1.1 \pm 0.3) 10^{-2} E^{-(1.1 \pm 0.1)}$$

photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$

corresponding to a reduced $\chi^2 \approx 1.1$ with 23 d.f. These two fits are shown in Fig. 2. During our observation, the power output in the range 0.5–5.0 MeV was the highest observed for NGC4151 ($\sim 2 \times 10^{45} \text{ erg s}^{-1}$ for a distance of 20 Mpc, $H_0 = 50 \text{ km s}^{-1} \text{Mpc}^{-1}$). In the adjacent decades, 50–500 keV and 5–50 MeV, the power was $\sim 2 \times 10^{44}$ and $\sim 6 \times 10^{44} \text{ erg s}^{-1}$ respectively.

Several models have been proposed to explain the emission from active galaxies up to γ -ray energies. The suggested mechanisms are the inverse Compton scattering of the synchrotron radiation photons^{17–20}, the accretion onto massive black holes^{21–23} or the rotation of magnetised objects (spinar model)²⁴. For the non-thermal models, the e-folding energy of the exponential fit (or the cutoff of the power fit) may arise from an equivalent cut-off in the electron spectrum that produces the γ -ray emission or may be due to the opacity of the source for pair production. For a thermal model interpretation, our data suggest a temperature of about 10^{10} K in the active region of the galaxy. An alternative may be the "Penrose Compton Scattering" model proposed by Leiter and Kafatos²⁵ in the hypothesis of the existence of a massive Kerr black hole at the centre of active galaxies. It is clear that models which predict a cutoff around 500 keV (see ref. 22) are not favoured by our spectral data.

Received 9 August; accepted 8 October 1979.

- Gursky, H., Kellogg, E. M., Leong, C., Tananbaum, H. & Giacconi, R. *Astrophys. J. Lett.* **165**, L43 (1971).
- Baity, W. A., Jones, T. W., Wheaton, Wm. A. & Peterson, L. E. *Astrophys. J. Lett.* **199**, L5 (1975).
- Ives, J. C., Sandford, P. W. & Penston, M. V. *Astrophys. J. Lett.* **207**, L159 (1976).
- Paciesas, W. S., Mushotzky, R. F. & Pelling, R. M. *Mon. Not. R. astr. Soc.* **178**, 23P (1977).

- Auriemma, G. et al. *Astrophys. J. Lett.* **221**, L7 (1978).
- Di Cocco, G. et al. *Nature* **270**, 319 (1977).
- Baker, R. E. et al. *Nucl. Instrum. Meth.* **158**, 595 (1979).
- Dolan, J. F. *Astrophys. Space Sci.* **17**, 472 (1972).
- Barr, P., White, N. E., Sanford, P. W. & Ives, J. C. *Mon. Not. R. astr. Soc.* **181**, 43P (1977).
- Mushotzky, R. F., Holt, S. S., & Serlemitsos, P. J. *Astrophys. J. Lett.* **225**, L115 (1978).
- Schonfelder, V. *Nature* **274**, 344 (1978).
- Schonfelder, V. Announcement of Kyoto, 16th ICRC (1979).
- Bignami, G. F., Fichtel, C. E., Hartman, R. C. & Thompson, D. J. preprint (1979).
- Elvis, M. *Mon. Not. R. astr. Soc.* **177**, 7P (1976).
- Meegan, C. A. & Haymes, R. C. preprint (1979).
- Zanrosso, E. M., Long, J. L., Zych, A. D., Gibbons D. & White, R. S. *Proc. 16th ICRC*, Kyoto, og 4–3 (1979).
- Bergeron, J. & Salpeter, E. E. *Astrophys. Lett.* **9**, 121 (1971).
- Bergeron, J. & Salpeter, E. E. *Astr. Astrophys.* **22**, 385 (1973).
- Grindlay, J. E. *Astrophys. J.* **199**, 49 (1975).
- Pinkau, K. *Proc. 16th ICRC*, Kyoto, og 4–16 (1979).
- Fabian, A. C., Maccagni, D., Rees, M. J. & Stoeger, W. R. *Nature* **260**, 683 (1976).
- Lightman, A. P., Giacconi, R. & Tananbaum, H. *Astrophys. J.* **224**, 375 (1978).
- Maraschi, L., Perola, G. C., Reina, C. & Treves, A. *Astrophys. J.* (in the press).
- Pacini, F. & Salvati, M. *Astrophys. J. Lett.* **225**, L99 (1978).
- Leiter, D. & Kafatos, M. *Proc. La Jolla Institute, UCSD-Workshop on Particle Acceleration Mechanisms on Astrophysics* (1979).

Evidence for low-energy γ -ray emission close to CG135+1

A. Della Ventura*, F. Perotti*, G. Villa*, G. Di Cocco†, R. C. Butler‡, A. J. Dean‡ & R. I. Hayles‡

* Laboratorio di Fisica Cosmica e Tecnologie Relative, CNR, Via Bassini, 15/A, Milan, Italy

† Laboratorio TESRE, CNR, Via De' Castagnoli 1, Bologna, Italy

‡ University of Southampton, Department of Physics, Southampton, UK

The COS-B source CG135+1 (ref. 1) is an extremely interesting object because of its possible identification with the QSO 0241+622. Radio observations have identified a strong point source within 2 arc s (ref. 2) of QSO 0241+622. X-ray emissions have also been observed from this region of the sky by the Uhuru (ref. 3), SAS-3 (ref. 2) and HEAO-A (ref. 4) telescopes. The HEAO-A (A1) data indicate that there are two soft X-ray sources in this region of the sky, H0241+62 which contains the QSO, and H0235+60 which includes the radio source GT0236+61 (ref. 5). The OSO-7 data in the energy range 1–40 keV (ref. 6) and measurements using the high-energy X-ray experiment on Ariel-V between 0.26 and 1.2 MeV (ref. 7) have indicated that the photon emission from this region of the sky is of an extremely hard nature, with a power law spectral index of $\gamma \approx 1.0$. However, the poor angular resolution of these telescopes make it uncertain whether the high-energy X-ray fluxes are due to the same celestial object. We now report the results of a balloon flight carried out on 8 October 1978 from Palestine, Texas, in which the region of the sky including the COS-B source CG135+1 was scanned by the Milan/Southampton telescope (MISO telescope). A 5σ excess in the counting rate of the telescope in the energy range 0.15–20 MeV was observed. The origin of this low-energy γ -ray emission is compatible with the measured position of CG135+1.

The MISO telescope used to study this region of the sky in the low-energy γ -ray region of the spectrum has been described elsewhere⁸ and consists of two scintillators, S1 (liquid scintillator) and S2 (sodium iodide), which form the central Compton coincidence detection system. The instrument was sensitive over the energy range 0.15–20 MeV and had an aperture of 3° FWHM in both azimuth and zenith planes for this observation. The telescope was stabilised in both the azimuth and zenith planes with a precision of 20 arc min by means of a two-axis orientation system.

Six drift scans were carried out to survey the region of the sky contained within the coordinate points 2 h 16 min and 3 h 12 min in right ascension and centred on 62° in declination, over the period 6–10 h in universal time. To obtain a reasonable number of scans of the source region during the 4-h observational period available, the telescope was driven in azimuth against the rotation of the sky with a speed of approximately 0.2 deg min^{-1} . The region of the celestial sphere studied encompasses the error boxes of the X- and γ -ray emissions measured by the Uhuru, HEAO-A, SAS-3, OSO-7, Ariel-V and COS-B telescopes. The measurements were made at a mean residual atmospheric pressure of 3.6 mbar with the telescope set at zenith angles in the range 30 – 34° . The results reported here refer to a preliminary analysis of the integral γ -ray counting rate above 150 KeV, for five passes. Each of the five drift scan measurements was analysed separately so as to eliminate effectively any systematic variations due to zenith and azimuth changes of the background counting rates. The housekeeping data and the spectra from γ -ray calibration sources were studied to ensure that there were no overall measurable changes in the gain of the system during this period. The only source of measurable systematic variation in the background counting rate was due to the change of altitude between 3.5 and 3.9 mbar, and the data have been corrected for this effect.

The background counting rate as a function of pressure was evaluated separately for each drift scan and for each of the three independent counting rate channels (S1, S2, S1.S2). After subtraction of the background a correlation was made between the residual counting rate and the angular distance between the axis of the telescope and the direction of CG135+1. For each pass there was an indication of an excess from the general

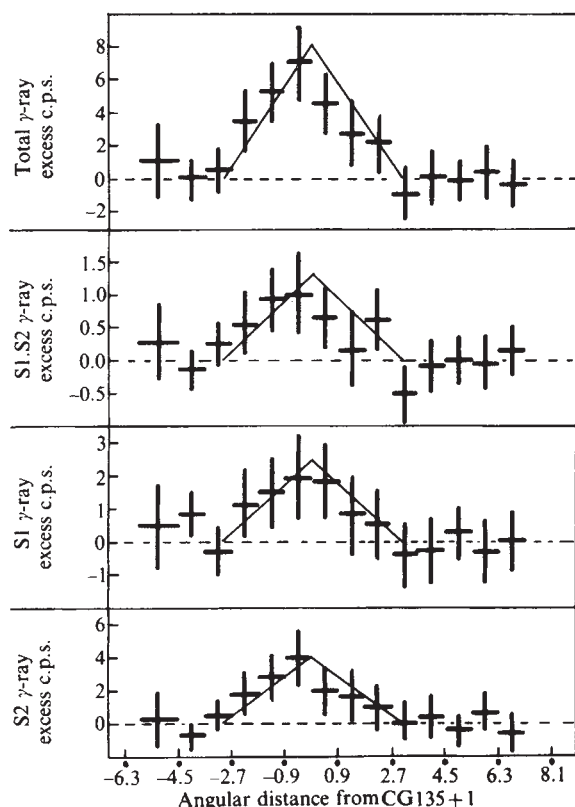


Fig. 1 The γ -ray excess plotted against the angular distance from CG135+1 is shown for the 'singles', the coincidence and the total events. The solid line represents the exposure of the centre of the COS-B error-box defined by the aperture of the MISO telescope (3° FWHM).

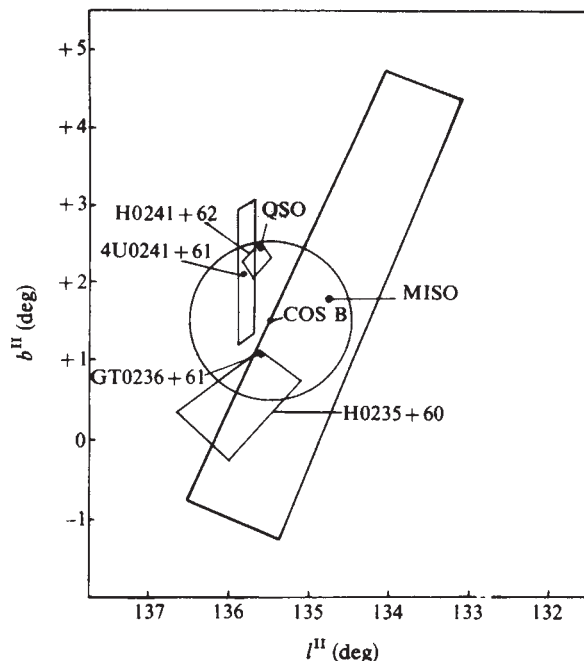


Fig. 2 The MISO error box in galactic coordinates (l^{II} = longitude; b^{II} = latitude) is shown together with other astronomical data. The MISO error box has 68% probability across the scan direction and 100% in the other.

direction of CG135+1. The results were combined statistically and are shown in Fig. 1. An excess is observed for each of the three independent channels (S1.S2 = 0.65 ± 0.2 c.p.s., S1 = 1.3 ± 0.5 c.p.s., S2 = 2.1 ± 0.6 c.p.s.). The statistical errors quoted here include the uncertainty in the evaluation of the variation of counting rate with residual pressure. Figure 1 also shows the data for the sum of all three channels in which an excess of 5σ is obtained (4.1 ± 0.8 c.p.s.). The solid line in Fig. 1 represents the exposure of the centre of the COS-B error box defined by the aperture of the MISO telescope (3° FWHM). The angular profiles of the excesses can be seen to follow closely the aperture of the MISO telescope.

Although the data indicate a 0.6° difference between the centroids of the MISO and COS-B emissions, the results are compatible with a common celestial origin. Figure 2 shows the MISO error box in galactic coordinates for the observed γ -ray flux together with other astronomical data taken at different photon energies. As our drift scans were made in one direction only, the limits of our error box are represented by the 1σ points along the direction of scan and by the total aperture of the telescope in the orthogonal direction (100% probability).

If it is assumed that the X-ray and COS-B γ -ray flux measurements originate from the same source object, a steepening of the spectrum is required above 1 MeV. On the basis of our integral counting rate of 4.1 ± 0.8 c.p.s. the spectral change must be close to 1 MeV. Only the full spectral analysis of our data will enable us to say more on the existence of such a change and consequently on the coincidence of the X- and γ -ray emission.

Received 7 August; accepted 3 October 1979.

1. Hermesen, W. *et al.* *Nature* **269**, 494 (1977).
2. Apparao, K. M. V. *et al.* *Nature* **273**, 450 (1978).
3. Forman, W. *et al.* *CFA preprint No. 763* (1977).
4. Share, G. M. *et al.* *Proc. XXI Cospar Meet. Innsbruck* (1978).
5. Gregory, P. C. *et al.* *Nature* **273**, 704 (1978).
6. Maraschi, L. *et al.* *Nature* **272**, 679 (1979).
7. Coe, M. J. *et al.* *Nature* **279**, 343 (1978).
8. Baker, R. E. *et al.* *Nucl. Instrum. Meth.* **158**, 595 (1979).

Separation of ^{26}Al and ^{26}Mg isobars by negative ion mass spectrometry

L. R. Kilius, R. P. Beukens, K. H. Chang, H. W. Lee & A. E. Litherland

Department of Physics, University of Toronto, Toronto, Canada M5S 1A7

D. Elmore, R. Ferraro & H. E. Gove

Nuclear Structure Research Laboratory, University of Rochester, Rochester, New York 14627

Aluminium-26 (half life 7.16×10^5 yr) is one of the important radioisotopes generated from the spallation by cosmic rays of argon in the Earth's atmosphere. It has been used in studies¹ of the growth rate of manganese nodules from the ocean floor and the constancy of the cosmic-ray intensity in the past². At present, the concentration of ^{26}Al ($^{26}\text{Al}/\text{Al} \leq 10^{-14}$) is measured in large samples by counting the γ rays following the β decay to ^{26}Mg . We have now demonstrated that it is possible to separate ^{26}Al and ^{26}Mg atoms in milligramme samples with a sensitivity for the ratio $^{26}\text{Al}/\text{Al}$ better than 10^{-14} , using negative ions, a tandem accelerator as a molecular disintegrator, and a mass spectrometer³. The measurement of $^{26}\text{Al}/^{10}\text{Be}$ ratios using the same apparatus and small samples, to provide a several million year chronology^{4,5} for oceanic sediments, now seems possible.

The complete separation of the isobars ^{26}Mg and ^{26}Al using negative ion mass spectrometry was shown to be possible in principle by Heinemeier and Hvelplund⁶ during measurements on the production of Mg^- and Al^- in sodium vapour. They observed that for 20 keV ions at equilibrium in sodium vapour the ratio of Al^-/Al^+ was 0.10 whereas the ratio of Mg^-/Mg^+ was less than 10^{-6} , implying a highly unstable magnesium negative ion.

Soon after the first detection by atom counting of ^{14}C (half life 5,730 yr) at natural abundances⁷, we tried unsuccessfully to use negative ions to detect ^{26}Al at levels of $^{26}\text{Al}/\text{Al}$ near 10^{-12} . It was realised during several of the early experiments that the problem was due to background from molecular fragments from ions such as $^{12}\text{C}^{14}\text{N}^-$, and from ^{27}Al positive ions from the injection of $^{27}\text{Al}^-$ ions. The abundant ^{27}Al positive ions were present because the mass resolution of the magnet at the ion source was not high enough to reject $^{27}\text{Al}^-$ ions when $^{26}\text{Al}^-$ ions were injected into the tandem accelerator. These problems can be solved by the addition of an electrostatic analyser, which specifies the ratio of the ion energy and the charge (E/q), to the magnetic analyser after the accelerator which specifies $(M/q) \times (E/q)$. M is the mass of the ions. The use of the combination of electrostatic and magnetic analysis is most successful when M and q do not contain common factors⁸.

After the installation of a 10° electrostatic analyser, successful measurements of $^{36}\text{Cl}/\text{Cl}$ ratios down to 10^{-15} (^{36}Cl half life 3.08×10^5 yr) were made and reported⁸. The apparatus has since been used⁹ to measure $^{10}\text{Be}/\text{Be}$ ratios to below 10^{-14} (^{10}Be half life 1.6×10^6 yr) and $^{36}\text{Cl}/\text{Cl}$ ratios in meteoritic and other samples¹⁰ to near 10^{-16} . The same apparatus⁸ has now been used to detect ^{26}Al . Recently, Raisbeck *et al.*¹¹ have reported the detection of ^{26}Al by accelerating positive ions to a high energy in a cyclotron and removing all the electrons so that $^{26}\text{Al}^{13+}$ and $^{26}\text{Mg}^{12+}$ can be completely separated in a magnetic field. A sensitivity for the ratio $^{26}\text{Al}/\text{Al}$ of 10^{-12} was reached. The complete stripping of the ions solved the problem¹² of detecting ^{26}Al with a cyclotron.

Two samples of aluminium and one of magnesium were prepared for use in the negative ion sputter source of the University of Rochester MP tandem Van de Graaff accelerator.

The first sample, of 99.999% pure aluminium, was used for background measurements and the second was machined from an aluminium-magnesium (1%) alloy which had been exposed some years ago to high energy electrons and photons in the beam dump of the 40 MeV Electron Linac at Toronto. The $^{26}\text{Al}/\text{Al}$ ratio was estimated to be less than 2×10^{-11} by measurements of the γ rays following the β decay of ^{26}Al in a 10-g sample. The magnesium sample was used to increase the sensitivity of the detection system for $^{26}\text{Mg}^-$ ions.

The negative ion current of $^{27}\text{Al}^-$ from the ion source was about 15 nA during the measurements. This is a low value compared with the current of $1 \mu\text{A}$ which has been reported from a negative ion sputter source (R. Middleton, personal communication) but it was found to be adequate for the demonstration of the separation of ^{26}Al and ^{26}Mg by negative ion mass spectrometry.

The negative ^{26}Al ions were accelerated to a total energy of 48.17 MeV in the tandem accelerator and the $^{26}\text{Al}^{5+}$ ions were analysed by the equipment used previously for ^{36}Cl detection. The dissociation of the molecular ions after charge exchange in the tandem accelerator and the electrostatic analysis provide an efficient way of eliminating such mass 26 molecules as $^{13}\text{C}_2$, $^{10}\text{B}^{16}\text{O}$, and $^{12}\text{C}^{14}\text{N}$ without the need for high mass resolution. Figure 1 illustrates in a two-parameter plot the clear separation between $^{26}\text{Al}^{5+}$ and $^{26}\text{Mg}^{5+}$ from the irradiated sample even though there was 1% magnesium in the sample. The measured $^{26}\text{Al}^{5+}$ counting rate of (1.5 ± 0.3) c.p.m. and the average analysed $^{27}\text{Al}^{5+}$ ion current of 5.3 nA implies a ratio for $^{26}\text{Al}/\text{Al}$ of $(3.8 \pm 0.7) \times 10^{-12}$ in the irradiated sample. During experiments with the unirradiated pure aluminium sample, no ^{26}Al counts were recorded in a total of 230 min; one count would have implied that an $^{26}\text{Al}/\text{Al}$ ratio of 10^{-14} was reached. To guard against drifting of the magnets, ion source fluctuations, and other systematic effects, the irradiated aluminium sample was counted immediately before and immediately after the background run on the pure Al sample. The counting rate of the irradiated sample was the same within statistical error for both of these runs, indicating that the system parameters had not changed significantly during the 230 min.

An estimate of the mass of aluminium sputtered during the experiment (1.39 mg) yielded an efficiency for producing negative ions Al^-/Al of about 10^{-3} and a total efficiency for detecting

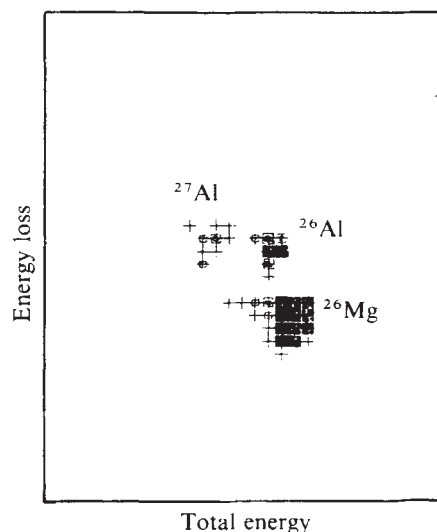


Fig. 1 Counts versus energy loss and total energy from an ionisation counter for an aluminium-magnesium (1%) alloy enriched in ^{26}Al . The few ^{27}Al counts detected at 46.4 MeV are some of the ions from multiple charge exchange in the accelerator tube which gives the required magnetic rigidity. A total of 34 counts corresponding to 48.17 MeV $^{26}\text{Al}^{5+}$ and 1,250 counts for $^{26}\text{Mg}^{5+}$ were recorded in 23 min. The ^{26}Mg contribution could be reduced substantially by chemical processing of natural samples. No ^{26}Al counts were observed from a pure aluminium sample.

$^{26}\text{Al}^{5+}/\text{Al}$ from the sample of about 10^{-4} . Even this low efficiency implies that the negative ion method is already superior to the γ -ray counting method for determining $^{26}\text{Al}/\text{Al}$ ratios below 10^{-12} in samples smaller than about 1 g.

The small current of Al^- ions used for these experiments can be increased by using a more powerful ion source. The efficiency for utilisation of the ^{26}Al atoms and the Al^- current can also be increased by starting with positive sputtered ions and charge exchanging in sodium vapour⁶. In either case currents over 1 μA should make possible measurements of $^{26}\text{Al}/\text{Al}$ ratios below 10^{-15} in under an hour from samples of aluminium possibly as small as 1 mg. This amount of aluminium can be found in as little as 10 mg of oceanic sediment⁴. The large tandem accelerator used for these demonstration experiments can also, in principle, be replaced by a much smaller tandem accelerator such as those at present under construction specifically for ^{14}C dating¹³ as the difference in energy loss (dE/dx) for ^{26}Al and ^{26}Mg , 10.7% in the present experiment, changes to only 7.1% for the smaller machine.

In Fig. 1, a peak corresponding to $^{26}\text{Mg}^{5+}$ can also be seen. It was demonstrated that the flux of $^{26}\text{Mg}^{5+}$ ions clearly peaked at the terminal voltage used in these measurements and also reached another much larger peak when the tandem terminal voltage was raised 58 ± 8 kV. These $^{26}\text{Mg}^{5+}$ ions from the larger peak are from the injection of $^{26}\text{MgH}^-$ molecular ions into the accelerator due to the inadequate resolution of the injection magnet ($M/\Delta M \sim 25$) or the breakup of $^{26}\text{MgH}_2^-$ into $^{26}\text{MgH}^-$ before magnetic analysis. The addition of an electrostatic analyser at the ion source would remove the second source of background. It is worth noting here that the aluminium alloy sample, containing 1% magnesium, enhances the yield of magnesium ions over that of a pure sample.

The $^{26}\text{Mg}^{5+}$ ions that appear at the same terminal voltage as the $^{26}\text{Al}^{5+}$ ions could be due to the injection of $^{26}\text{Mg}^-$ into the accelerator. This, we believe, is an unlikely explanation as the very low limit⁶ for the formation of Mg^- can best be explained by a mean lifetime for Mg^- of less than about 10^{-7} s. As the acceleration time for the negative ions in the present experiment is about 10^{-5} s, no Mg^- ions should survive the acceleration. In addition, the efficiency for formation of such metastable ions in a negative sputter ion source is expected to be very low. (R. Middleton, personal communication). The absence of Mg^- ions in the observations of Heinemeier and Hvelplund⁶ could also be due to the low binding energy of the electron in Mg^- and to the destruction¹⁴ of Mg^- in the electric analysing field. If this occurs, it can be estimated¹⁴ that the electron binding energy of the negative ion would have to be less than 1 meV. A Mg^- ion with such a low binding energy would be destroyed completely in the high electric fields of the tandem accelerator. These fragile $^{26}\text{Mg}^-$ ions could also contribute to the observed $^{26}\text{Mg}^{5+}$ ions as a result of multiple charge exchange processes. Further research into the origin of this $^{26}\text{Mg}^{5+}$ peak is desirable although it is not expected to interfere with the detection of ^{26}Al in aluminium provided the magnesium abundance is less than 1%.

We thank K. H. Purser for useful discussions and the General Ionex corporation for the loan of the 10° electrostatic analyser. We also thank G. M. Raisbeck and R. Middleton for communicating their unpublished results. The work was supported by the NSERC of Canada and the USNSF. R.P.B. is a Killam Fellow.

Received 13 August; accepted 11 October 1979.

- Guichard, F., Reyss, J. L. & Yokoyama, U. *Science* **272**, 155–156 (1978).
- Hampel, W. & Schaeffer, O. A. *Earth planet. Sci. Lett.* **42**, 348–358 (1978).
- Purser, K. H., Litherland, A. E. & Gove, H. E. *Nuclear Instr. Meth.* **162**, 637–656 (1979).
- Lal, D. *J. Oceanogr. Soc. Japan, 20th Anniv. Vol.*, 600–614 (1962).
- Reyss, J. L. *et al. Science* **193**, 1119–1120 (1979).
- Heinemeier, J. & Hvelplund, P. *Nuclear Instr. Meth.* **148**, 425–429 (1978).
- Purser, K. H. *et al. Rev. Phys. Appl.* **12**, 1487–1492 (1977).
- Elmore, D. *et al. Nature* **277**, 22–25; 246 (1979).
- Kilius, L. R. *et al. Bull. Am. phys. Soc.* **24**, 650 (1979).
- K. Nishizumi, K. *et al. Earth planet. Sci. Lett.* (in the press).
- Raisbeck G. M. *Le Journal de Phys. Lett.* **40**, 241–244 (1979).
- Müller, R. *Science* **196**, 489–494 (1977).
- Purser, K. H. & Hanley, P. R. *Proc. 1st Conf. on Radiocarbon Dating with Accelerators* (ed. Gove, H. E.) 165–168 (Rochester, 1978).
- Oparin, V. A. *et al. Sov. Phys. Lett.* **13**, 249–252 (1971).

A maximum speed of wetting

T. D. Blake & K. J. Ruschak

Research Laboratories, Eastman Kodak Company, Rochester, New York 14650

The wetting (or dewetting) of a solid by a liquid is an integral part of many important processes such as coating, petroleum recovery, distillation and the handling of liquid fuels in low gravity conditions. Several experiments^{1–4} have shown that wetting lines (where liquid, air and solid phases meet) which are straight at slow rates of movement over the solid have a sawtooth shape at sufficiently high speeds. We now offer a quantitative explanation for this phenomenon based on the postulate that, for a given system, there is a maximum rate at which wetting can proceed. The consequences of this interpretation are likely to be important, since, in many practical situations, the aim is to maximise the speed of wetting without entraining the displaced phase.

In our experiments, a tape plunges vertically into a pool of liquid. We measure the dynamic contact angle (the angle that the liquid/air interface makes with the moving tape), and we observe the behaviour of the wetting line. At low tape speeds, the wetting line is straight and horizontal. As the speed is increased, the wetting line is pulled below the level of the free liquid surface and the advancing contact angle measured through the liquid increases, reaching 180° at some critical speed U . At tape speeds exceeding U , the wetting line breaks into two or more straight-line segments inclined to the horizontal at an angle ϕ which increases steadily with speed. In other words, the wetting line adopts a sawtooth configuration (Fig. 1). At sufficiently high speeds, air bubbles break off the trailing (deeper) vertices and are entrained into the liquid.

Wetting has been modelled⁵ as an adsorption/desorption process in which, as the wetting line advances, molecules adsorbed at the initial solid/liquid interface are randomly but steadily displaced by molecules from the advancing fluid and a new, more or less equilibrium solid/fluid interface is created. This model has been analysed^{5,6} using the statistical mechanical theory of transport processes developed by Eyring and co-workers⁷, and an expression has been derived relating the dynamic contact angle to the speed of wetting. The form of the

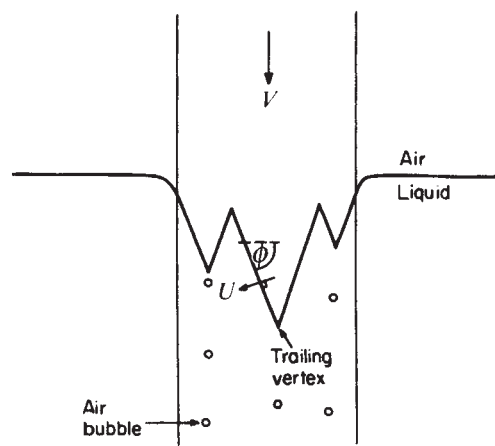


Fig. 1 Schematic representation of the sawtooth wetting line obtained when a tape plunges vertically into a pool of liquid at a speed, V , greater than the maximum speed of wetting, U . The wetting line shows both leading and trailing vertices, with air bubbles entrained at the latter.

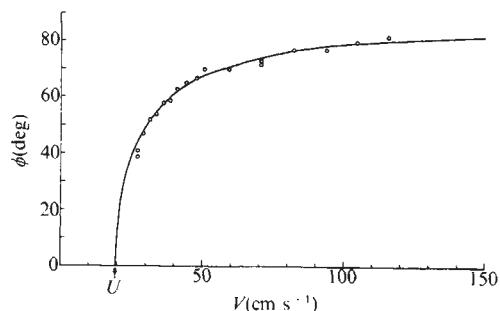


Fig. 2 Measured values of ϕ , the angle of inclination of the wetting line with respect to the horizontal, for a polyester tape (Kodak Estar film base) 5 cm in width and 0.018 cm in thickness plunging at speed V into a nominally 85.6% glycerol/water mixture (viscosity 1.0 P, surface tension 66 dyn cm^{-1}). The theoretical curve was fitted to the data to determine the maximum speed of wetting, $U = 19.4 \text{ cm s}^{-1}$.

relationship leads to the interesting prediction that there is a maximum speed at which a wetting line can advance normal to itself, that is, a maximum rate of wetting per unit length of wetting line (area/time). This maximum corresponds to a dynamic advancing contact angle of 180° . Conversely, there is a maximum rate of dewetting corresponding to a dynamic receding contact angle of 0° . According to the argument set out below, the breakup of the wetting line into segments slanted from the horizontal is evidence of these speed limitations.

If we suppose that there is a maximum speed, U , for wetting corresponding to a dynamic contact angle of 180° , and if we further suppose that it is the component of tape velocity V normal to the wetting line which is relevant to the wetting

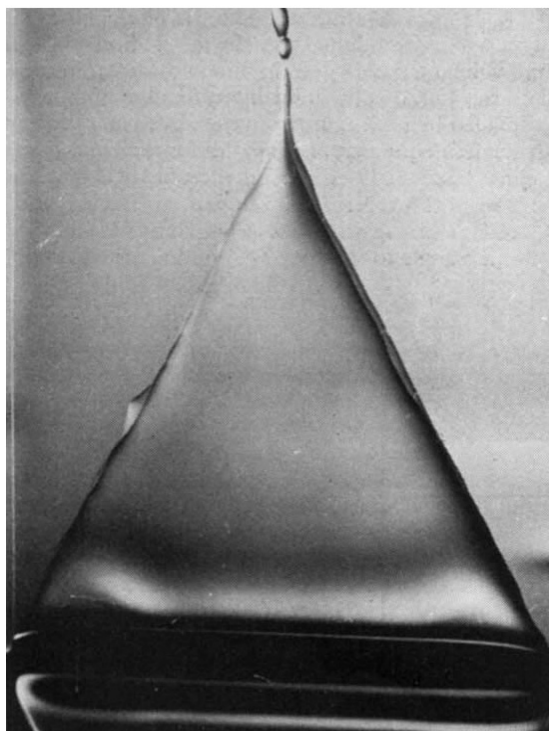


Fig. 3 Example of the sawtooth wetting line that forms when the tape is slowly withdrawn from the liquid at a steady speed greater than the maximum speed of dewetting. The tape is transparent so that two nearly identical triangular tongues of liquid can be distinguished, one on either side of the tape. Methylene blue dye was added to the liquid to increase contrast for the photograph.

process and not the tape velocity itself, then it follows that the tape can be wet at speeds exceeding U only if the wetting line slants from the horizontal at an angle, ϵ , which ensures that the component of tape velocity normal to the wetting line does not exceed U . That is, it follows that

$$\cos \epsilon \leq U/V; \quad (V \geq U)$$

To test this relationship, we have measured ϕ as a function of speed for a polyester tape plunging vertically into an aqueous glycerol solution. The data are shown in Fig. 2. The curve through the data is the limit of the above inequality, and so it appears that the wetting line segments adopt the minimum possible inclination such that

$$\cos \epsilon = U/V; \quad (V \geq U)$$

One consequence of this is that the minimum length of wetting line necessary to wet a unit width of the tape is

$$l = V/U; \quad (V \geq U)$$

and we see that at tape speeds greater than U , the length of the wetting line must increase linearly with tape speed. The value of U obtained by fitting the data in Fig. 2 is 19.4 cm s^{-1} . At this speed, the measured contact angle is 170° rather than 180° , but the agreement with theory is reasonable in the light of experimental uncertainties. The speed at which air bubbles begin to be entrained is 25 cm s^{-1} .

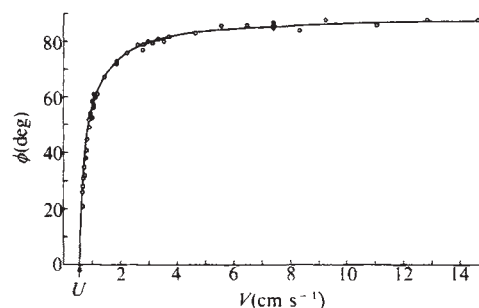


Fig. 4 Measured values of ϕ , the angle of inclination with respect to the horizontal, when the tape is withdrawn from a nominally 72% glycerol/water mixture (viscosity 0.23 P). The theoretical curve was fitted to the data to determine the maximum speed of dewetting, $U = 0.55 \text{ cm s}^{-1}$.

The same phenomenon occurs when the tape is slowly withdrawn from the pool and the liquid recedes from the tape. Above a tape speed which presumably corresponds to a dynamic contact angle of 0° , the wetting line again consists of at least two straight segments slanted from the horizontal (Fig. 3), and drops or a continuous rivulet of liquid may be entrained from the trailing (higher) vertices. Measured values of the wetting line inclination for one system are shown in Fig. 4. Again, the curve is the limit of the above inequality, where U is the maximum speed of dewetting. The value of U obtained by fitting the data is 0.55 cm s^{-1} , and liquid drops begin to be entrained into the air phase at a speed of 0.8 cm s^{-1} .

The theory presented above predicts that the breakup of the wetting line will be postponed to tape speeds greater than U when the tape does not enter the liquid vertically. Specifically, the wetting line will adopt a sawtooth shape once the tape speed exceeds $U/\cos \psi$, where ψ is the angle of the tape with respect to the vertical. Preliminary experiments in which the tape entered or left the pool at a 45° angle have shown that the expected gain is in fact realised.

The theory also suggests why, with a tape plunging vertically into a pool at a speed greater than U , air bubbles form only at the

trailing vertices of the sawtooth wetting line. Presumably, the curvature of the wetting line is large but finite where two straight-line segments seem to intersect. Thus, the tangent to the wetting line at this apparent vertex is horizontal, and this point of the wetting line will be drawn into the liquid since its speed relative to the tape cannot exceed U . On the other hand, the leading vertices will tend to recede and become trailing vertices, accounting perhaps for the unsteadiness of the sawtooth wetting line when leading vertices are present.

Received 20 July; accepted 11 October 1979.

1. Deryagin, B. V. & Levi, S. M. *Film Coating Theory* 140–143 (Focal, London, 1964).
2. Perry, R. T. thesis, Univ. Minnesota (1967).
3. Burley, R. & Kennedy, B. S. *Chem. Engng Sci.* **31**, 901–911 (1976).
4. Wilkinson, W. L. *Chem. Engng Sci.* **30**, 1227–1230 (1975).
5. Blake, T. D. thesis, Univ. Bristol (1968).
6. Blake, T. D. & Haynes, J. M. *J. Colloid Interface Sci.* **30**, 421–423 (1969).
7. Glasstone, S., Laidler, K. J. & Eyring, H. *The Theory of Rate Processes* (McGraw-Hill, New York, 1941).

A CO₂-climate sensitivity study with a mathematical model of the global climate

Syukuro Manabe & Ronald J. Stouffer

Geophysical Fluid Dynamics Laboratory/NOAA,
Princeton University, Princeton, New Jersey 08540

An increase in the CO₂-content of the atmosphere resulting from man's activity could have a significant effect on the climate in the near future¹. We describe here some new results from a study of the response of a mathematical model of the climate to an increase in the CO₂-content of the air.

The mathematical model consists of (1) a general circulation model of the atmosphere and (2) a simple mixed layer ocean model with uniform thickness. The atmospheric model predicts the changes of the vertical component of vorticity, divergence, temperature, moisture and surface pressure based on the equations of motion, the thermodynamical equation, and the continuity equations of moisture and mass. The horizontal distributions of these variables are represented by a limited number of spherical harmonics (maximum zonal wavenumber retained is 15)^{2,3}. However, the vertical distributions are specified at nine unequally spaced finite difference levels. The model has a global computational domain and realistic geography.

For the computation of solar and terrestrial radiation, the distributions of ozone and cloud cover are prescribed beforehand and the concentration of carbon dioxide is set differently for each experiment, whereas the distribution of water vapour is determined from the prognostic system of water vapour.

Condensation of water vapour is predicted whenever supersaturation is indicated in the computation of the continuity equation of water vapour⁴. Snowfall is predicted when air temperature near the Earth's surface falls below the freezing temperature⁵. Otherwise, rainfall is predicted.

The temperature of continental surface is determined so that it satisfies the requirement of the heat balance⁵. The changes of soil moisture and snow depths are obtained from the budget-computations of water and snow, respectively⁵.

The ocean model is a static isothermal water layer of uniform thickness with provision for a sea ice layer. The thickness of 68 m is chosen to ensure that the heat storage associated with the annual cycle of observed sea surface temperature is correctly modelled. Ocean temperature change is computed based on the budget among surface heat fluxes. In the presence of sea ice, the temperature of underlying water is at the freezing point and the

heat flux through ice is balanced by the latent heat of freezing and melting at the bottom of the ice. This process, together with the melting at the ice top, sublimation and snowfall, determines the change of ice thickness⁶. The albedo of sea ice and continental snow is assumed to vary between 60 and 70% depending on latitudes. Smaller values are assigned for thin sea ice or thin snow or melting sea ice.

To evaluate the climatic effect from an atmospheric CO₂ increase, we perform two time integrations of the model assuming the normal seasonal insolation cycle and starting from an isothermal, dry and motionless atmosphere with isothermal ocean. The atmospheric CO₂ concentration is set at 300 p.p.m. and 1,200 p.p.m. by volume, respectively (hereafter, these experiments are referred to as 1×CO₂ and 4×CO₂ experiments). Both cases settle down to a stable climatic condition in about a decade of model time (see Fig. 1). The standard experiment successfully reproduces the observed basic characteristics of geographical and seasonal temperature variation, encouraging the assertion of realism for the model sensitivity to CO₂ changes. The seasonal variations of the model climate discussed below represent mean annual cycles over the last 3-yr periods of both time integrations.

Figure 2a shows the seasonal variation of the difference of zonal mean surface air temperature between the 4×CO₂ and 1×CO₂ atmospheres. In low latitudes, the warming due to the quadrupling of CO₂ content in air is relatively small and depends little on season, whereas in high latitudes, it is generally larger and varies markedly with season particularly in the Northern Hemisphere. Over the Arctic Ocean and its neighbourhood, the warming is at a maximum in early winter and is small in summer. This implies that the range of seasonal variation of surface air temperature in these regions reduces significantly in response to the quadrupling of CO₂ content of air. From the seasonal variations of zonal mean sea ice thickness shown in Fig. 3, it is observed that the sea ice from the 4×CO₂ experiment is everywhere less than the sea ice from the 1×CO₂ experiment. Therefore it is reasonable to suggest that the 1×CO₂ atmosphere is insulated by thicker sea ice from the influence of offlying seawater and has a more continental climate with a larger seasonal variation of temperature than the 4×CO₂ atmosphere. Although the poleward retreat of highly reflective snowcover and sea ice is mainly responsible for the relatively large warming

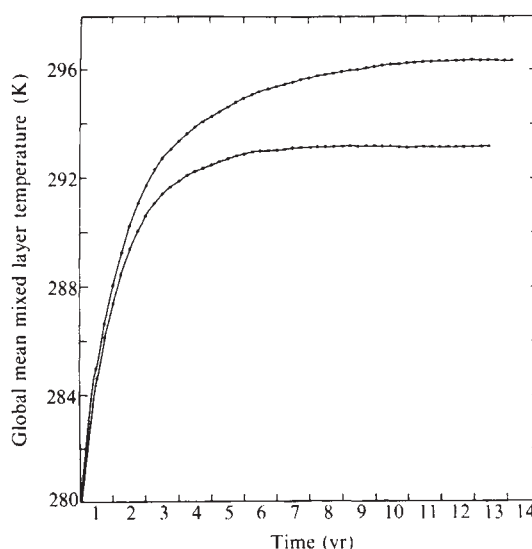


Fig. 1 Time variation of the global mean mixed layer temperature. Lower and upper lines are obtained from the 1×CO₂ and 4×CO₂ experiment, respectively. The seasonal variation is removed through the application of the running mean procedure over the period of 12 months.

in high latitudes, the change of the thermal insulation effect of sea ice strongly influences the seasonal variation of the warming over the Arctic region.

An analysis of the heat balance over the Arctic Ocean indicates that, in early winter, the flux of sensible heat from the Earth's surface to the atmosphere is much larger in the $4 \times \text{CO}_2$ case than the $1 \times \text{CO}_2$ case because of the difference in the magnitude of upward conductive heat flux through sea ice. As the stable stratification in the surface layer of the model atmosphere prevents the sensible heat flux from penetrating into the middle troposphere in high latitudes, the aforementioned difference in sensible heat flux results in the large warming of surface air during early winter. As previously stated, the magnitude of the warming in the summer is much less than the warming in the winter. Because sea ice is thin or absent during summer, the surface albedo reduces significantly and net incoming solar radiation increases from the $1 \times \text{CO}_2$ to the $4 \times \text{CO}_2$ case. However, the additional solar radiation is used either for melting upper surface of sea ice or warming the ice-free mixed layer which has a large heat capacity. Thus, the summer warming of the surface air turns out to be relatively small. However, the additional solar energy, which is absorbed during summer in the

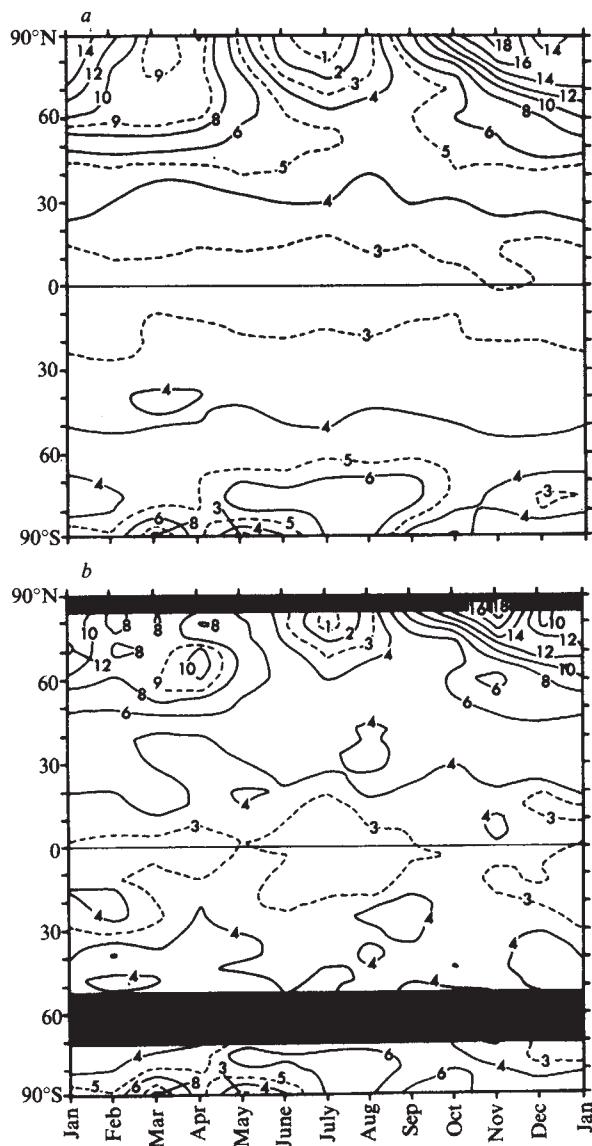


Fig. 2 Latitude-time distributions of the difference in zonal mean surface air temperature between the $4 \times \text{CO}_2$ and $1 \times \text{CO}_2$ atmospheres. *a*, Oceans and continents; *b*, continents only. (Positive value indicates warming caused by the quadrupling of CO_2 -content of air.)

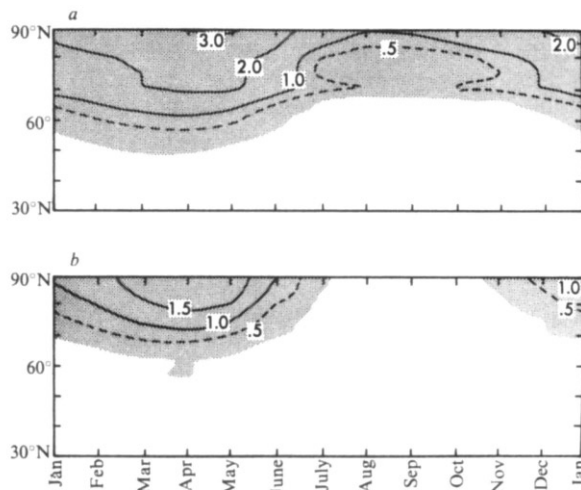


Fig. 3 Latitude-time distributions of zonal mean thickness of sea ice in metres for the Northern Hemisphere oceans. *a*, $1 \times \text{CO}_2$ experiment; *b*, $4 \times \text{CO}_2$ experiment. The shading indicates the areas where sea ice thickness is more than 0.1 m.

$4 \times \text{CO}_2$ case, delays the appearance of sea ice or reduces its thickness. This increases the conductive heat flux through ice in early winter when the air-sea temperature difference becomes large, thereby enhancing the warming of the surface atmospheric layer in early winter.

The seasonal variation of the difference in the surface air temperature between the two experiments over the model continents is significantly different from the variation over the model oceans. According to Fig. 2*b*, which shows the latitude-time distribution of the difference in zonal mean surface air temperature over continents, the CO_2 -warming in high latitudes is at a maximum in early winter, being influenced by the large warming over the Arctic Ocean discussed above. However, Fig. 2*b* also indicates a secondary centre of relatively large warming around 65°N in April. This results from a large reduction in surface albedo in spring when the insolation acquires a near-maximum intensity. The reduction of snow cover area from the $1 \times \text{CO}_2$ to the $4 \times \text{CO}_2$ experiment is responsible for this albedo difference.

This study shows two interacting mechanisms, each acting to produce its own sensitivity maximum. The maximum warming of the early winter over the Arctic Ocean and its neighbourhood is caused by the change in sea ice thickness, and the relatively large warming over the continents in spring is produced by snow albedo feedback. Therefore, it seems reasonable that simpler climate models with the albedo feedback mechanism but without the thermal insulation effect of sea ice would have the largest sensitivity in the early summer^{7,8}. Using his simple model, Sellers⁹ found an early winter maximum sensitivity at $\sim 80^\circ\text{N}$. This winter maximum, however, apparently results from different mechanisms.

Table 1 summarises the area mean changes of the annual mean surface air temperature of the model atmosphere which occurs in response to the quadrupling of CO_2 content in air. According to Table 1, the global mean warming of the model

Table 1 The surface air temperature difference between the $4 \times \text{CO}_2$ and $1 \times \text{CO}_2$ atmosphere in $^\circ\text{C}$

	Northern Hemisphere	Southern Hemisphere	Global domain
Difference	4.5	3.6	4.1

Positive value indicates warming caused by the quadrupling of CO_2 -content of air.

atmosphere is $\sim 4^{\circ}\text{C}$. This result suggests that the warming caused by the doubling of CO_2 -content would be $\sim 2^{\circ}\text{C}$. This is significantly less than the warming which is estimated by the general circulation model of Manabe and Wetherald¹⁰ with idealised geography and without seasonal variation of insolation. Table 1 also reveals that the area mean warming of the Northern Hemisphere is significantly larger than that of the Southern Hemisphere. This interhemispheric difference results partly from the smallness of the snow albedo-feedback effect over the Antarctic ice sheet where surface albedo differs little between the $1 \times \text{CO}_2$ and $4 \times \text{CO}_2$ -experiments.

In view of the assumption of fixed cloudiness and various simplifications contained in the sea ice modelling, the quantitative aspect of the present results should be received with caution. However, this study suggests that the warming of the atmosphere in response to an increase in CO_2 -content of the air will have significant seasonal and interhemispheric asymmetries.

Received 6 August; accepted 19 September 1979.

1. *Energy and Climate* (National Academy of Sciences, Washington, D.C., 1977).
2. Bourke, W., McAvaney, B., Puri, K. & Thurling, R. *Meth. comput. Phys.* **17**, 267–324 (1977).
3. Gordon, T. & Stern, B. *The GARP Programme on Numerical Experimentation*, Rep. No. 7, 46–87 (1974).
4. Manabe, S., Smagorinsky, J. & Strickler, R. F. *Mon. Weath. Rev.* **93**, 769–798 (1965).
5. Manabe, S. *Mon. Weath. Rev.* **97**, 739–774 (1969).
6. Bryan, K. *Mon. Weath. Rev.* **97**, 806–827 (1969).
7. Held, I. M. *Rep. of the JOC/SCOR Joint Study Conf. on General Circulation Models of the Ocean and Their Relation to Climate*, V53–V65 (ICSU and WMO, Geneva, 1977).
8. Ramanathan, V., Lian, M. S. & Cess, R. D. *J. geophys. Res.* **84**, 4949–4958 (1979).
9. Sellers, W. D. *J. appl. Met.* **13**, 831–833 (1974).
10. Manabe, S. & Wetherald, R. T. *J. atmos. Sci.* **32**, 3–15 (1975).

Carbon disulphide and carbonyl sulphide from biogenic sources and their contributions to the global sulphur cycle

V. P. Aneja, J. H. Overton & L. T. Cupitt

Environmental Sciences Center, Northrop Services Incorporated,
Research Triangle Park, North Carolina 27709

J. L. Durham & W. E. Wilson

US Environmental Protection Agency, Research Triangle Park,
North Carolina 27711

Estimates of the magnitude of biogenic sulphur emissions range from $\sim 70\%$ of the total atmospheric sulphur burden^{1–7}, and the chemical nature of the emissions has not been clearly established. Conway¹ speculated that the principal volatile biogenic component of the sulphur cycle was hydrogen sulphide (H_2S) whereas Lovelock *et al.*⁸ and Rasmussen⁹ have suggested that dimethyl sulphide (DMS) contributes to the apparent source deficits. Aneja *et al.*¹⁰ have shown that both H_2S and DMS are emitted from saltmarshes. Other volatile sulphur compounds which may contribute to the sulphur burden of the atmosphere include methyl mercaptan (CH_3SH)^{11,12}, dimethyl disulphide ($(\text{CH}_3)_2\text{S}_2$)^{11,12}, carbonyl sulphide (COS)¹³ and carbon disulphide (CS_2)^{14,15}. We report here the discovery of CS_2 and COS emanating from a saltmarsh and estimate their emission rates using emission flux reactor and bag chamber techniques. The species CS_2 and COS are relatively inert in the troposphere, so may be assumed to penetrate to the stratosphere, where they may be photolysed to form the sulphur dioxide (SO_2) and sulphates (SO_4^{2-}) known to be present in the stratosphere. Based on the measured fluxes, we show that the emissions from marshes are important to the sulphate aerosol burden ($\leq 19\%$) of the stratosphere, but not important for the tropospheric sulphur burden ($\leq 0.2\%$).

An emission flux reactor (chamber) was used for measuring Earth-atmosphere fluxes of biogenic sulphur compounds. The chamber technique has high sensitivity for flux measurement without the necessity of measuring very low concentrations. In addition, gas residence times in the emission flux reactor are of the order of minutes so that chemical transformations between emission and analysis may be minimised. The chamber used by Aneja¹¹ and Hill *et al.*¹² was modified by making the walls from 5-mil (~ 0.13 mm) thick FEP Teflon supported by an exterior aluminium frame. This design was chosen to provide negligible attenuation of ambient light.

The analytical instrument used was a gas chromatograph (Tracor 270HA) equipped with a flame photometric detector. A 394-nm interference filter made the detector sensitive to sulphur compounds. A 36-foot (~ 11 m) FEP Teflon column ($\frac{1}{8}$ inch (0.32 cm) o.d.) packed with 40–60 mesh Teflon coated with 5% polyphenyl ether (PPE) and 0.5% phosphoric acid was used at 50°C to separate gaseous sulphur species¹⁷. The carrier gas was nitrogen with a flow rate of ~ 81 ml min^{-1} . The retention time for CS_2 in these conditions is ~ 14 min which allows a good separation from $\text{H}_2\text{S} + \text{COS}$ (1.5 min), SO_2 (3.1 min), and CH_3SH (5.0 min). The overlapping DMS (12.2 min) and CS_2 peaks were manually resolved into separate contributions. This column did not resolve H_2S and COS , so these species were separated by incorporating a second 19-foot (5.8 m) FEP Teflon column ($\frac{1}{8}$ inch (0.32 cm) o.d.) packed with 50/80 mesh Porapak QS in parallel with the first. The two columns were incorporated in the flow lines separately by a pneumatically controlled switching system. The carrier gas for the second column was nitrogen with a flow rate of ~ 40 ml min^{-1} . The retention time for COS in these conditions is ~ 8 min, which allows a good separation from H_2S (retention time ~ 5.5 min); however, this column retained CS_2 . The gas chromatograph was calibrated for both columns in the laboratory and in the field using a dilution system and bag samples.

Intensive measurements were made for biogenic sulphur species at Cox's Landing, Long Beach, North Carolina; and at Cedar Island Wildlife Refuge, North Carolina, which are saltmarshes on the east coast of the US (Fig. 1). Experiments were performed during late summer at both locations. At Cox's Landing, 12 different test sites were selected. At each test site, two identical emission flux reactors were placed next to each other. One was flushed with pure gaseous nitrogen (de-aerated) and the other with ambient air (aerated). This was done to gain information on the oxygen-depleted gaseous environment, as well as to maintain the integrity of the emitted sulphur compounds in the de-aerated case. Sampling was done over mud flats, over regions close to marsh grass where the vegetation was decaying, and over marsh grass (*Spartina alterniflora*) clipped to about 1 inch above ground level. Time-integrated bag samples of the effluent gas from the emission flux reactor were collected every 30 min. These bag samples were analysed within 30 min for various gaseous sulphur species in a controlled environment mobile laboratory parked at the edge of the marsh. All the experiments at Cox's landing were performed to study the emission fluxes of CS_2 for diurnal and tidal variations. During the experiments, the ambient temperature ranged between 19.1 and 33.4°C , sediment temperature with the tide out ranged from 24.2 to 32.6°C , and the water temperature with the tide in was between 23.8 and 31.2°C .

Figure 2a is a typical chromatogram (using the PPE column) of sulphur gases emitted from the saltmarsh. It indicates the presence and identity of CS_2 and various other biogenic sulphur species.

The average emission rates of CS_2 at the various types of test sites are given in Table 1. In 20 of the 172 aerated runs, CS_2 was detected; and CS_2 appeared only one in the de-aerated case. All of the data in Table 1 are for the aerated case. The CS_2 emission rates in the mud flat zone in the presence and absence of a water column for varying light conditions were less than 0.05 g m^{-2} yr^{-1} (this is the emission rate based on the lowest detectable CS_2 concentration). CS_2 is also evolved over and in

Fig. 1 *a*, Location of the sites. Sites 1–12 were at Cox's Landing and Sites 13–15 at Cedar Island.

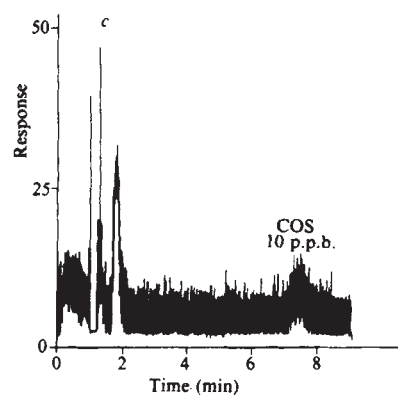
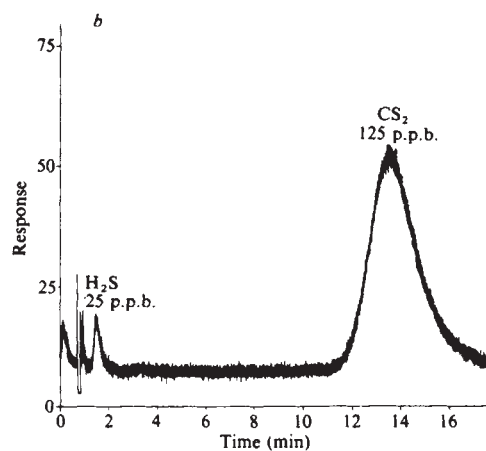
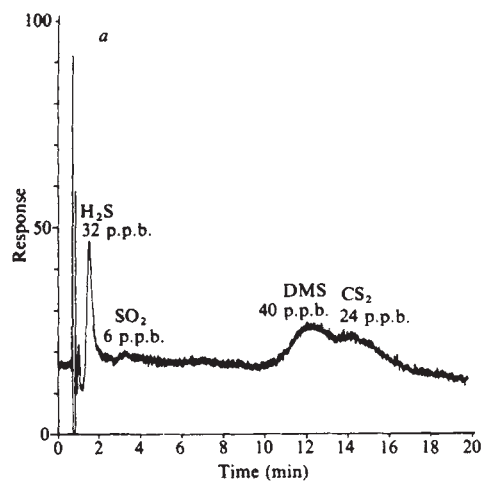
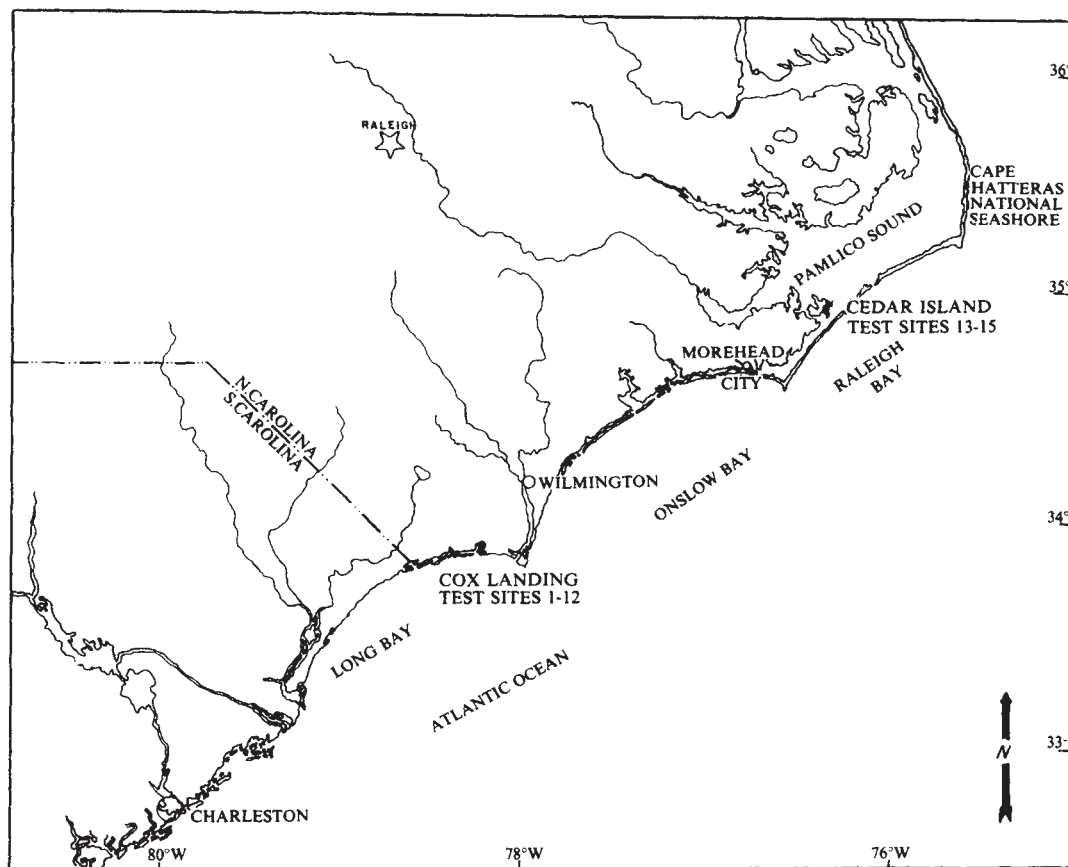


Fig. 2 *a*, Chromatogram of volatile sulphur compounds from *Spartina* zone. The chromatogram was obtained at Cox's Landing using the PPE column and a nonlinear sulphur-specific flame photometric detector (FPD). *b*, Chromatogram of volatile sulphur compounds obtained from enclosed *Spartina* stem. The chromatogram was obtained at Cox's Landing using the PPE column and a nonlinear sulphur-specific FPD. *c*, Chromatogram of volatile sulphur compounds resolved on a Porapak QS column. The sample was obtained at Cedar Island.

Table 1 Emission rate measurements of CS₂

Test site characteristic	Mean emission rate of all runs, gS m ⁻² yr ⁻¹ (no. of runs)	Mean of runs with emission rate greater than detectable limit*†, gS m ⁻² yr ⁻¹ (no. of runs)
Over mud flat zone (direct emission rate measurement*) August 1977	<0.05 (100)	— (0)
Over decaying <i>Spartina</i> zone (direct emission rate measurement*) August 1977	~0.04 (61)	0.19 ± 0.07 (13)
Over <i>Spartina alterniflora</i> zone (direct emission rate measurement*) August 1977	0.13 (11)	0.21 ± 0.06 (7)
Enclosed single <i>Spartina alterniflora</i> plant (indirect emission rate measurement†) July 1978	0.20 (10)	0.20 ± 0.09 (10)

* Direct emission rate $n = (F \cdot \Delta C / A)$, where F = steady volumetric flow rate of gas through the chamber; ΔC = concentration increase; A = area of emitting surface covered by the chamber. Based on the CS₂ detectable limit of the instrument, the corresponding emission rate is <0.05 gS m⁻² yr⁻¹, and the CS₂ analytical error is ±3% of the estimated value.

† Indirect emission rate $n = (S \cdot \Delta C \cdot V) / t \times 8,760$, where S = no. of stems per m²; ΔC = concentration increase, parts per 10⁹; V = volume of chamber (1.96 × 10⁻³ m³); t = time of enclosure (h). For this measurement, the CS₂ detectable limit is 0.017 gS m⁻² yr⁻¹ and the CS₂ analytical error is ±3% of the estimated value.

‡ Values are given ±s.d.

the vicinity of *Spartina alterniflora* under the aerated emission flux reactor in the presence and absence of a water column and the fluxes range from <0.05 to ~0.3 gS m⁻² yr⁻¹ (Table 1). CS₂ may be formed at a juncture of the carbon and sulphur cycles by either biogenic or chemical activity. These observations show that CS₂ is certainly produced in an aerated environment, and that it may be formed in a de-aerated environment. Lovelock¹⁵ has observed that CS₂ can originate in anaerobic conditions on the sea floor, while, in the *Spartina alterniflora* zone, we observe it more often in aerobic conditions. The reasons for the differences between the aerated and de-aerated chamber runs are not yet known.

These observations suggest that *Spartina alterniflora* plants may be associated with the production of CS₂. This association is also supported by the results of an indirect method of estimating gaseous sulphur flux (see ref. 18 for the enclosure method). A single stem was enclosed for ~6 h in a Pyrex glass cylindrical chamber of ~1.96 litre volume (~100 cm long and ~5 cm i.d.). Samples were collected during day and night. Ten such experiments were performed and *in situ* syringe samples (one for each stem) were withdrawn from the glass chambers and analysed on the gas chromatograph. In all cases large quantities of CS₂ were observed (see Fig. 2b and Table 1).

Seneca *et al.*¹⁹ have determined the stem density in North Carolina coastal saltmarshes to be ~200 ± 25 stems m⁻². Based on this result, the emission rate of CS₂ from *Spartina alterniflora* stems is estimated to be ~0.2 gS m⁻² yr⁻¹ (Table 1). Note that this indirect emission rate measurement leads to emission rate values similar to those obtained by direct flux measurement over the *Spartina* zone (Table 1).

At Cedar Island three different test sites were selected over *Spartina alterniflora* and mud flat zones. At each test site the emission flux reactor was used in the aerated mode only. All the experiments described (Table 2) were performed to study the emission fluxes of COS. It establishes the presence and identity of COS as indicated by the gas chromatogram in Fig. 2c. The indirect method of estimating gaseous sulphur flux revealed no release of COS from *Spartina alterniflora* plants.

To estimate the per cent of biogenic contribution of CS₂ and COS to the global tropospheric sulphur cycle, we assumed that all marshes are completely made up of *Spartina alterniflora* and that they emit uniformly. We took the total marsh area²⁰ as 3.8 × 10⁵ km², and assumed that 100 × 10⁶ tons of gaseous sulphur are emitted per year from the biosphere by natural processes⁴. Based on our CS₂ indirect and COS direct emission rates (~0.2 gS m⁻² yr⁻¹ and 0.03 gS m⁻² yr⁻¹, respectively) and the above assumptions, the contribution of CS₂ and COS from marshes is a small portion, <0.07% and 0.09%, respectively, of the biogenic sulphur. Obviously, more work is needed to characterise the global processes adequately.

We also estimated the contributions of CS₂ and COS emissions to the stratospheric sulphate aerosol layer. According to Crutzen's¹⁶ estimate, a flux of ~1.8 × 10⁻⁴ gS m⁻² yr⁻¹ into the stratosphere will account for the entire stratospheric sulphate aerosol layer. Applying Junge's²¹ one-dimensional eddy diffusion model to the troposphere, and using the present CS₂ and COS average emission rate for global marshes and CS₂ and COS atmospheric lifetimes¹⁴ of 1 yr and 20 yr respectively, an upper limit to the contributions can be made by assuming zero CS₂ and COS concentrations at 15 km height (average Junge aerosol layer ~20 km). The upper limit contributions are 8% for CS₂ and 11% for COS. However, if land plants emit CS₂ at rates comparable to those measured for *Spartina alterniflora* the biogenic emissions of CS₂ could account for the entire Junge layer. Experiments are under way to test for CS₂ emissions for land vegetation.

We thank E. Corse, R. Whitkus, M. Richards, and H. Sher of NSI for assistance.

Table 2 Emission rate measurements of COS

Test site characteristic	Mean emission rate of all runs, gS m ⁻² yr ⁻¹ (no. of runs)	Mean of runs with emission rate greater than detectable limit‡, gS m ⁻² yr ⁻¹ (no. of runs)
Over mud flat zone (direct emission rate measurement*) July 1978	0.05 (2)	0.05 ± 0.01 (2)
Over decaying <i>Spartina</i> zone (direct emission rate measurement*) July 1978	0.01 (2)	0.02 (1)
Over <i>Spartina alterniflora</i> zone (direct emission rate measurement*) July 1978	0.03 (7)	0.04 ± 0.03 (6)
Enclosed single <i>Spartina alterniflora</i> plant (indirect emission rate measurement†) July 1978	<0.01* (10)	<0.01* (0)

* Based on the COS detectable limit of the instrument, the corresponding emission rate is <0.01 gS m⁻² yr⁻¹ and the COS analytical error is ±20% of the estimated value. See Table 1.

† See Table 1.

‡ Values given ±s.d.

Received 21 June; accepted 27 September 1979.

- Conway, E. J. *Proc. R. Ir. Acad. A* **48**, 119–159 (1943).
- Eriksson, E. J. *geophys. Res.* **68**, 4001–4008 (1963).
- Junge, C. E. *Air Chemistry and Radioactivity* (Academic, New York, 1963).
- Robinson, E. & Robbins, R. C. *Sources, Abundance, and Fate of Gaseous Atmospheric Pollutants: SRI Project Report PR-6755* (American Petroleum Institute, New York, 1968).
- Kellogg, W. W., Cadle, R. D., Allen, E. R., Lazrus, A. L. & Martell, E. A. *Sciences* **175**, 587–596 (1972).
- Friend, J. P. in *Chemistry of the Lower Atmosphere* (ed. Rasool, S. I.) 177–201 (Plenum, New York, 1973).
- Granat, L., Hallberg, R. O. & Rodhe, H. *Ecol. Bull. (Stockholm)* **22**, 39–134 (1976).
- Lovelock, J. E., Maggs, R. J. & Rasmussen, R. A. *Nature* **237**, 452–3 (1972).
- Rasmussen, R. A. *Tellus* **26**, 254–260 (1974).
- Aneja, V. P., Overton, J. H. Jr, Cupitt, L. T., Durham, J. L. & Wilson, W. E. *Tellus* **31**, 174–178 (1979).
- Aneja, V. P. thesis North Carolina State Univ. Raleigh (1975).
- Hill, F. B., Aneja, V. P. & Felder, R. M. *Envir. Sci. Hlth* **13**, 199–225 (1978).
- Hanst, P. L., Spiller, L. L., Watts, D. M., Spence, J. W. & Miller, M. F. *J. Air Pollut. Control Ass.* **25**, 1220–1226 (1975).
- Sandalls, F. J. & Penkett, S. A. *Atmos. Envir.* **11**, 197–199 (1977).
- Lovelock, J. E. *Nature* **248**, 625–626 (1974).
- Crutzen, P. J. *Geophys. Res. Lett.* **3**, 73–76 (1976).
- Stevens, R. K., Mulik, J. D., O'Keefe, A. E. & Krost, K. J. *Analyt. Chem.* **43**, 827 (1971).
- Zimmerman, P. R. *Testing of Hydrocarbon Emissions from Vegetation and Development of a Methodology for Estimating Emission Rates from Foliage* (EPA Contract No. DU-77-C063, Research Triangle Park, 1978).
- Seneca, E. D., Stroud, L. M., Blum, U. & Noggle, G. K. *An Analysis of the Effects of the Brunswick Nuclear Power Plant on the Productivity of Spartina alterniflora (Smooth Cordgrass) in the Dutchman Creek, Oak Island, Snow's Marsh, and Walden Creek Marshes, Brunswick County, North Carolina, 1975–1976* (Third Annual Report to Carolina Power and Light Company, Raleigh, 1976).
- Woodwell, G. M., Rich, P. M. & Hall, C. A. S. *A.E.C. Symp. Ser.* **30**, 221–240 (1973).
- Junge, C. *Proc. int. Conf. of Structure, Composition and General Circulation of the Upper and Lower Atmospheres and Possible Anthropogenic Perturbations* (Melbourne, 1974).

Tidal flexure of ice shelves measured by tiltmeter

S. N. Stephenson & C. S. M. Doake

British Antarctic Survey, Natural Environment Research Council,
Madingley Road, Cambridge, UK

J. A. C. Horsfall

Department of Geodesy and Geophysics, University of Cambridge,
Madingley Rise, Cambridge, UK

A new design of tiltmeter has been used to measure the tidal bending of the Ronne Ice Shelf, Antarctica. By this means the point at which the ice sheet leaves the land and starts floating on the sea has been defined to within 600 m. Monitoring this position in future should give an indication of the state of balance of the West Antarctic ice sheet, often considered¹ to be unstable even if not actually disintegrating. The hydrostatic level meter is light and simple and can operate unattended while recording continuously for periods of several weeks. Observations at eight different sites on Rutford Ice Stream (Fig. 1) gave tidal records on ice more than 1,700 m thick and at points up to 600 km from the nearest open sea.

Antarctic ice shelves are floating ice sheets varying in thickness from 100 m to more than 1,700 m. They rise and fall on the ocean tide, bending over a narrow region where they join the grounded ice sheet. The variation in theoretical surface elevation profiles over the duration of a tidal cycle have been derived^{1,2} for several constitutive laws, but there have been no published measurements of tidal changes in profiles. The reason is partly because there are crevasse fields in many hinge zones and partly because of the lack of a suitable measurement technique. Gravimeters, strain meters and optical methods all have serious drawbacks because of their cost or their logistic and manpower requirements. The hydrostatic level meter, however, is a stable, continuously recording and relatively cheap instrument that, once placed in a shallow trench in the snow surface to reduce temperature and atmospheric effects, can operate

unattended for long periods. For a short base length (5 m in our case) the instrument will measure changes in surface gradient which can be integrated along a transect to obtain a surface profile.

There are several reasons for wanting to measure tidal flexing. Actual profiles can be compared with theoretical^{1,2} profiles in order to deduce the most appropriate constitutive law. There is evidence that the crystal orientation fabric of the ice can be significantly altered as it passes across the hinge zone³. A knowledge of strain rates is necessary for calculating the development of the fabric⁴ and its effect on the flow law. Plastic deformation of ice in the bending zone dissipates tidal energy⁵, but for a precise estimate both the constitutive law and the region over which bending occurs must be known. By measuring tidal constants it will be possible to study the interaction of tide not only with the main body of an ice shelf⁶ but also with its grounding region. The stability of ice sheets grounded below sea level is thought to be controlled by the shielding effect of ice shelves⁷; tidal flexing may effect the dynamic behaviour in the hinge zone and may thus be an important consideration in the possible disintegration of the West Antarctic ice sheet⁸. By locating accurately the position of the grounding line any future changes can easily be monitored.

An instrument was required which would measure tilts of the ice surface as small as 10^{-6} radians with as large a dynamic range as possible. Tiltmeters currently available were considered unsuitable mainly because of the difficulty of mounting them on ice; all were intended to be installed in boreholes to give the necessary stability. In most materials and particularly in ice, such a hole takes several days to become mechanically and dimensionally stable. We therefore decided to use an instrument which could measure the height difference between two separated points on the glacier surface. The hydrostatic level meter consists of two vessels containing a liquid connected by a tube so that the liquid surface forms a gravitational potential (Fig. 2); a second tube connects the air space above the liquid at each end. Instruments of this form have been used for many purposes, for example for precision levelling on Hawaii⁹ but the liquid height has been read by direct optical means and attempts to make an automatically recording instrument using floats with displacement transducers, ultra-sonic measurement of fluid level, or mercury capacitance transducers have only been a success in laboratory conditions.

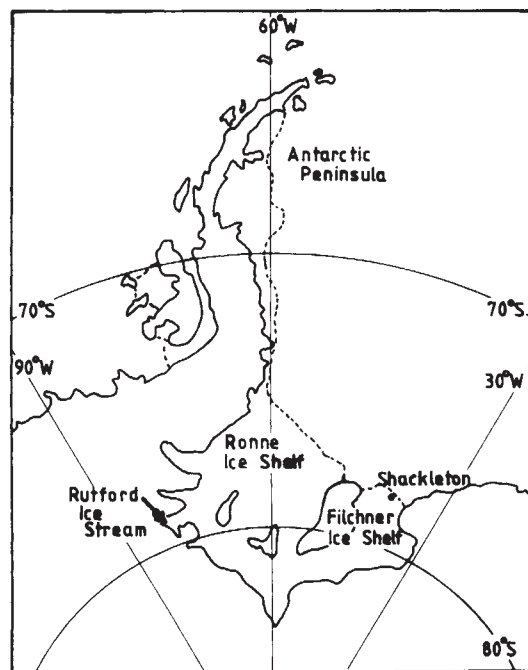


Fig. 1 Map showing location of tiltmeter sites.

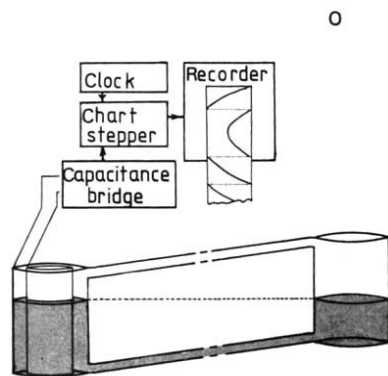


Fig. 2 Hydrostatic level meter with concentric cylinder transducer to measure liquid height.

To monitor liquid height the new instrument uses a sensitive capacitance bridge to measure the electrical capacitance between two concentric cylinders containing a liquid of known dielectric constant. Rather severe constraints on the electrical and physical behaviour of the liquid are all met by an inert fluorocarbon, ICI MFL 18. The liquid and air vent tubes are $\frac{1}{4}$ -inch diameter silicone rubber tubing which does not become stiff at low temperatures. The length of the instrument was chosen to be 5 m as a compromise between the height difference signal and ease of handling; the dimensions of the measuring capacitor give a total usable range of 100 mm representing a tilt of $\pm 10^{-2}$ radians. The electronic system is the same as that of Horsfall and King's geophysical tiltmeter¹⁰ and has a resolution of 10^{-6} in liquid height, representing tilts of 2×10^{-7} radians. It consists of a transducer drive unit, a chart stepper, a chart recorder and a quartz clock time unit; by expanding the range of the recorder by a factor of 25, the chart stepper allows full advantage to be taken of the wide range of the instrument. The complete system has a power consumption of less than 10 mA at 12 V and may therefore be run from either low temperature dry cells or, as in this experiment, from a vehicle battery.

Eight tiltmeter sites were chosen along a stake scheme which ran down the approximate centre line of the ice stream. The upper two sites showed no response to tidal bending while the third, 600 m downstream, and the rest all gave strong tidal records. The settling time until equilibrium was reached was small because the end units are light and because the instrument measures a height difference. At one site which showed no tidal movement it was possible to measure a drift which lasted for only 4 h. The shortest recording period at any of the sites which gave tidal records was 3 days. The amplitude of the vertical movement, about 10 mm, had not appreciably died away at the end site, showing that the flexure zone extended more than 6 km. The distribution of amplitudes did not show the pattern expected from considering a simple hinge line. This suggests that the glacier side walls influence the flexing, or that there are

locally grounded regions within the flexing zone downstream of the first grounding line. The measured curves can be compared with a tidal prediction (Fig. 3) based on data for the four main tidal constituents (O_1 , K_1 , M_2 , S_2) obtained at Shackleton Base on the Filchner Ice Shelf in 1957 by Pratt¹¹. Although the Rutford Ice Stream is nearly 1,000 km from Shackleton Base and differs in longitude by 43° , the times of high and low water match surprisingly well.

The ease by which tidal measurements can be taken using the new tiltmeter means that the interaction between ice shelves and tides can be studied in detail. In particular, our results show that the position of the grounding line can be found with sufficient accuracy for any changes to become quickly apparent when monitored in future.

We thank H. E. Thompson, R. V. Otto and J. L. W. Walton for help in collecting the data.

Received 20 August; accepted 15 October 1979.

1. Hughes, T. *Rev. Geophys. Space Phys.* **13**, 502–526 (1975); **15**, 1–46 (1977).
2. Holdsworth, G. *Ann. Geophys.* **33**, 133–146 (1977).
3. Woodruff, A. H. W. & Doake, C. S. M. *J. Glaciol.* **23**, 223–232 (1979).
4. Budd, W. F. *Z. Gletscher. Glaziol.* **8**, 65–105 (1972).
5. Doake, C. S. M. *Nature* **275**, 304–305 (1978).
6. Williams, R. T. & Robinson, E. S. *Science* **203**, 433–445 (1979).
7. Thomas, R. H. & Bentley, C. R. *Quat. Res.* **10**, 150–170 (1978).
8. Thomas, R. H., Sanderson, T. J. O. & Rose, K. E. *Nature* **277**, 355–358 (1979).
9. Eaton, J. P. *Bull. seism. Soc. Am.* **49**, 301–316 (1959).
10. Horsfall, J. A. C. & King, G. C. P. *Nature* **274**, 675–676 (1978).
11. Pratt, J. G. D. *Trans-Antarctic Expedition 1955–58. Sci. Rep. No. 4* (1960).

A geologically meaningless Rb–Sr total rock isochron

D. Field

Department of Geology, University of Nottingham, Nottingham, UK

A. Råheim

Mineralogisk-Geologisk Museum, Sars Gate 1, Oslo, Norway

A major objective in total rock Rb–Sr isotope studies of relatively homogeneous igneous or high-grade metamorphic gneisses is to obtain as wide a range of $^{87}\text{Rb}/^{86}\text{Sr}$ ratios as possible, so that the resultant isochron may be well defined¹. This is often achieved by collecting samples at some distance from each other. Recent studies^{2–6} have shown that primary Rb–Sr isotopic systems can be profoundly disturbed, and even completely reset, in response even to low-grade secondary deformations and/or mineral alterations, and in three cases^{2–4} isochron diagrams based on regionally collected samples have been shown to yield linear arrays corresponding to apparent ages intermediate between the true ages of the primary and secondary events. The evidence we present here shows that it is also possible for well defined intermediate isochron apparent ages to be produced where this sampling technique is used on rock suites which show evidence of only minor secondary alteration effects. An isochron so produced is geologically meaningless, that is, the apparent age is unrelated to any geological event.

We have recently presented the results of a detailed Rb–Sr isotope investigation of a high-grade, acid-intermediate, charnockitic orthogneiss suite at Arendal, southern Norway⁴. Eight series of samples, collected from individual outcrops, yielded apparent ages in two distinct groups, at $\sim 1,540$ Myr and $\sim 1,060$ Myr. The major conclusions were that the $\sim 1,540$ Myr age (best estimate $1,536 \pm 26$ Myr; $^{87}\text{Rb}/^{86}\text{Sr}$ initial ratio

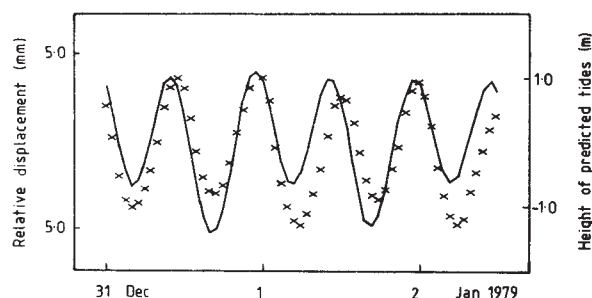


Fig. 3 Comparison between tiltmeter record (crosses) and predicted tide (solid line).

Table 1 Analytical data for regionally collected charnockite samples, Flotsøya, southern Norway

Sample	Rb	Sr	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
F178	228.1	89.5	7.4775	0.84638 ± 0.00014
F179	182.8	98.5	5.4161	0.81007 ± 0.00006
F180	181.8	90.4	5.8822	0.81789 ± 0.00010
F405	235.8	76.2	9.0755	0.87686 ± 0.00010
F406	256.2	83.8	8.9944	0.87021 ± 0.00014
F408	144.2	129.2	3.2500	0.77014 ± 0.00012
F409	150.5	104.4	4.2302	0.78683 ± 0.00012
F509	235.0	79.1	8.7302	0.87085 ± 0.00014

Rb, Sr and Rb/Sr determined by X-ray fluorescence⁹. Precision of $^{87}\text{Rb}/^{86}\text{Sr}$ is $\pm 1\%$. Precisions of $^{87}\text{Sr}/^{86}\text{Sr}$ quoted as 2σ standard error of mean. Details of location, petrography and analytical procedures presented previously⁴.

0.70345 ± 0.00014) reflects the formation of the high-grade charnockite mineralogy (quartz-plagioclase-orthopyroxene $\pm$ K-feldspar, hornblende, clinopyroxene, biotite, garnet) and that there was subsequent disturbance of the total rock isotopic systems during a low-grade event which occurred at the same time as the intrusion of undeformed granite sheets and pegmatite dikes at $1,063 \pm 20$ Myr. The only petrographically visible effects of this superimposed low-grade event are (1) partial, and often variable, alteration of orthopyroxenes to chlorite and/or serpentine, (2) minor corrosion of pre-existing biotite, and (3) incipient turbidity within some plagioclases. These mineral alterations were initiated along narrow (≤ 1 mm) irregularly spaced cracks which facilitated the open system behaviour. The superimposed event caused resetting to the younger age not only of all primary mineral systems investigated (plagioclase, K-feldspar, biotite, in four localities), but also of the 1–2 kg rock systems at those localities (three) where the mineral alterations are most pronounced. Both the secondary total rock ages, and all mineral ages, fall within error of the $1,063 \pm 20$ Myr age obtained for one of the intrusive granite sheets. For each of the localities which suffered resetting of the total rock systems, all samples with $^{87}\text{Rb}/^{86}\text{Sr} \geq 0.8$ were displaced below the original 1,536 Myr isochron. This showed that the resetting process did not involve a simple rehomogenisation of the Rb–Sr systems within individual outcrops, but that there was a systematic change in $^{87}\text{Rb}/^{86}\text{Sr}$ and/or $^{87}\text{Sr}/^{86}\text{Sr}$.

The results presented here (Table 1) refer to a group of regionally-collected samples from the same charnockite body. They have been taken from an area of 1 km^2 at the northern end of the island of Flotsøya⁴; the area centres on one of the previously described localities (no. 4) which has suffered isotopic resetting during the secondary event. Both series of data are plotted in Fig. 1, together with the mineral data for sample 4.4. The data for the samples from locality 4 (a 10-m road cut) were presented in the earlier study⁴.

The five total rock samples from locality 4 give a good fit isochron age of $1,075 \pm 122$ Myr with a value for the mean square of the weighted deviates (MSWD)⁷ of 1.10; the mineral (plagioclase, K-feldspar, biotite) + host rock data for sample 4.4 yield $1,036 \pm 20$ Myr (MSWD = 2.69). When the mineral data are regressed together with those for all of the rocks from this locality, an eight-point isochron age of $1,035 \pm 9$ Myr (MSWD = 1.19) is obtained (Fig. 1). These ages are all within error (2σ) of the $1,063 \pm 20$ Myr age for the intrusive granite sheet, and they correlate with the low-grade superimposed event. The regionally collected samples define a good-fit apparent age of $1,264 \pm 25$ Myr. An MSWD of 1.58 indicates that the scatter about the isochron can be assigned to experimental error. The problem is that this good-fit 1,264 Myr isochron apparent age correlates with neither the primary ($\sim 1,540$ Myr) nor the secondary ($\sim 1,060$ Myr) event, and there is no geological evidence for an intermediate event to which it might relate. We interpret the

$1,264 \pm 25$ Myr isochron 'age' as being geologically meaningless.

A key aspect of these data is that a combination of the locality 4 samples with the regional samples yields an apparent age of $1,259 \pm 24$ Myr (Fig. 1), which is virtually identical to that given by the regionally collected samples alone; the MSWD value, although slightly higher at 2.07, is still indicative of reasonable fit. The implication is obvious: in a regionally-based sampling programme, any individual sample collected from locality 4 would plot on, or very close to, the geologically meaningless isochron, despite the fact that the within-locality age is quite different. This situation arises because of the relatively small dispersion of $^{87}\text{Rb}/^{86}\text{Sr}$ at locality 4.

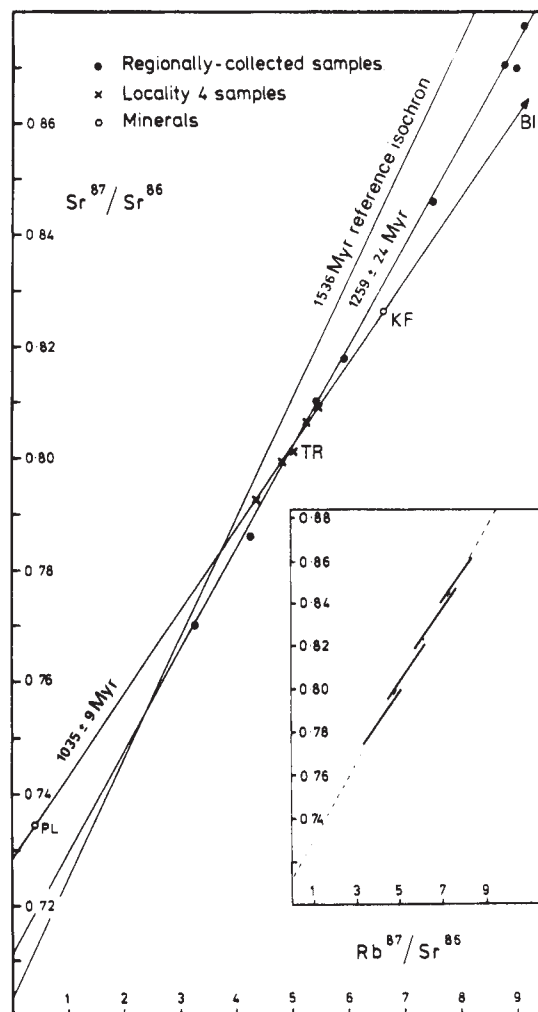


Fig. 1 Rb–Sr isochron diagram for regionally collected rock samples (●), and minerals (○) and rocks (×) from a single locality (no. 4). The $1,536 \pm 26$ Myr reference isochron ($\text{IR} = 0.70345 \pm 0.00014$) relates to the formation of the high grade charnockite mineralogy⁴. The $1,035 \pm 9$ Myr isochron age (MSWD = 1.19; $\text{IR} = 0.72862 \pm 0.00024$) is that given by a combination of rocks and minerals at locality 4. The $1,259 \pm 0.00024$ Myr apparent age (MSWD = 2.07; $\text{IR} = 0.71189 \pm 0.00182$) is that given by a combination of the regionally collected ($n = 8$) and locality 4 ($n = 5$) rock samples. By themselves the regionally collected samples define an apparent age of $1,264 \pm 25$ Myr (MSWD = 1.58; $\text{IR} = 0.71107 \pm 0.00194$). The data have been regressed by the method of York¹⁰. The decay constant used is $1.42 \times 10^{-11} \text{ yr}^{-1}$. All errors are quoted as $\pm 2\sigma$. The inset represents the probable relationship between the geologically meaningless 1,264 Myr isochron (dashed line) and within-locality sub-isochrons ($\sim 1,060$ Myr, solid lines). The latter would reflect the true age of the superimposed event.

All of the data points for the regionally-collected samples plot below the 1,536 Myr primary reference isochron (Fig. 1). Because (1) the displacement is in the same sense as that shown by samples with $^{87}\text{Rb}/^{86}\text{Sr} > 0.8$ from all of the localities (including locality 4) which are known to have equilibrated to the 1,060 Myr age, and (2) only the two apparent age groups ($\sim 1,540$ Myr and $\sim 1,060$ Myr) were revealed by the within-outcrop sampling⁴, we interpret the present positions of the regionally-collected samples as meaning that each was collected from a subvolume which suffered disturbance or resetting during the $\sim 1,060$ Myr superimposed event. If each of these isotopic systems was completely rehomogenised, the sub-volumes would actually describe a number of parallel sub-isochrons corresponding to the $\sim 1,060$ Myr event, as occurs at locality 4. Moreover, if it is assumed that each sub-volume would most likely be characterised by a limited range in $^{87}\text{Rb}/^{86}\text{Sr}$ such as occurs at locality 4 and elsewhere⁴, any sample taken from any locality would plot approximately along the 1,264 Myr, but geologically meaningless, apparent isochron. The model is schematically illustrated in the inset to Fig. 1. From this, it is evident that certain combinations of regionally collected samples would probably give rise to significant scatter, in which case it could at least be deduced that there had been some isotopic disturbance. More importantly, the data presented here show that regional sampling in this type of geological situation can also produce a combination which yields a good-fit intermediate apparent isochron (here, $1,264 \pm 25$ Myr). Despite a lack of significant scatter this isochron does not record a real isotopic event and is, therefore, geologically meaningless.

The model we have suggested is similar to a strontium isotopic equilibration model suggested by Roddick and Compston⁸, with the important difference that, in the present case, there must have been systematic changes in $^{87}\text{Rb}/^{86}\text{Sr}$ and/or $^{87}\text{Sr}/^{86}\text{Sr}$ at each locality in order to displace all of the samples below the original 1,536 Myr isochron. The sub-volumes cannot have remained closed with respect to the Rb-Sr isotopes during the 1,060 Myr event. We have argued elsewhere that the main factor was net loss of ^{87}Sr , rather than increases in the Rb/Sr ratios⁴.

In the example we have described, the $^{87}\text{Sr}/^{86}\text{Sr}$ initial ratio of the meaningless isochron (0.71107 ± 0.00194) is significantly higher than that given by the rocks which define the primary event (0.70345 ± 0.00014)⁴, but other evidence suggests that this is not always so and that significantly lower $^{87}\text{Sr}/^{86}\text{Sr}$ initial ratios may also be generated (data not shown).

So far, we have related the problem of geologically meaningless total rock isochrons only to regionally collected samples, but the same situation might well arise within single outcrops if there is locally a wide dispersion in the $^{87}\text{Rb}/^{86}\text{Sr}$ ratios and isotopic resetting was only achieved within smaller sub-volumes. In such cases, the only sure way to test the scale of homogenisation during resetting would be to undertake isotopic analysis of minerals and of several total rock samples from each of a number of sub-volumes within the outcrop.

Our final conclusion is the same as that recently deduced by Bell and Blenkinsop²: selection of samples for isochron-age analysis should not be made solely on the basis of seeking the maximum possible spread in the $^{87}\text{Rb}/^{86}\text{Sr}$ ratios.

This work was supported by NATO grant 1391. We thank T. Enger and J. Eyett for technical assistance.

Visual lateralisation effect in reading Chinese characters

Ovid J. L. Tzeng, Daisy L. Hung & Bill Cotton

University of California, Riverside, California 92521

William S.-Y. Wang

University of California, Berkeley, California 94720

Visual half-field experiments have been used to investigate perceptual differences in reading different types of writing scripts. It has been observed that tachistoscopic recognition of phonetic-based scripts (such as the English alphabet and Japanese Kana) tends to show a right visual field-left hemisphere (RVF-LH) superiority effect^{1,2} whereas recognition of logographic symbols (for example, Chinese characters) tends to show a left visual field-right hemisphere (LVF-RH) superiority effect^{3,4}. A cerebral orthography-specific localisation hypothesis has been proposed to account for these data. We have conducted three experiments to examine the visual lateralisation effect in reading logographic symbols such as Chinese characters. Chinese subjects in the first experiment were exposed to brief tachistoscopic presentation of a single character, and their task was to name the character as soon as possible. A LVF-RH superiority was found. In the second experiment, the stimuli were two vertically arranged characters and the subjects were asked to name the stimulus terms as soon as possible. The third experiment was the same as the second except that the task was to decide whether these character-strings were correct semantic terms with a manual response. A RVF-LH superiority effect was found in both the second and third experiments. These differential visual lateralisation results were interpreted as reflecting the function-specific property of the two hemispheres and cast doubt on the orthography-specific localisation hypothesis proposed by earlier investigators.

There has been much confusion in the literature as to whether asymmetries of function in the two sides of the brain arise from stimulus properties or from the nature of task requirements. Results of some recent clinical and experimental studies on the recognition of Japanese words have generally been used as evidence to support the former view. Among the many writing systems in the world today, Japanese is unique in that three types of script are used simultaneously to represent text. Most words in Japanese are written in Kanji, a logographic script, while grammatical particles are usually written in hiragana, and loan words (that is foreign words such as television) in katakana. The latter two are called Kana scripts and each symbol represents a syllable of the spoken language. So, a fluent reader of Japanese has to know all three types of script. Sasanuma and her associates⁵ have presented evidence that the ability of Japanese aphasic patients to use the Kanji and Kana scripts may be selectively related to type of aphasic disorder. Moreover, such differential properties can also be demonstrated in visual half-field experiments on normal subjects, with a LVF-RH superiority for the recognition of Kanji characters and a RVF-LH superiority for the Kana symbols. The latter result is similar to those obtained with alphabetic writings. This implies that the underlying reason for this difference is the requirement of visual-to-sound mapping. Although there is indeed evidence for the left hemisphere specialisation for speech sound⁶, generalisation of such findings to explain the differences between reading logographic symbols and reading alphabetic/syllabary symbols may be misleading. There is much evidence that reading logographic symbols also requires speech recoding⁷. Thus, what the hemispheric difference found in the tachistoscopic recognition of Kanji and Kana (or alphabetic) symbols reflects is not an

Received 11 June; accepted 2 October 1979.

1. Faure, G. *Principles of Isotope Geology* (Wiley, New York, 1977).
2. Bell, K. & Blenkinsop, J. *Nature* **273**, 532-534 (1978).
3. Page, R. W. J. *geol. Soc. Amer.* **25**, 141-164 (1978).
4. Field, D. & Råheim, A. *Earth planet. Sci. Lett.* **45**, 32-44 (1979).
5. Black, L. P. et al. *Tectonophysics* **54**, 103-137 (1979).
6. Hawkesworth, C. J. & Morrison, M. A. *Nature* **276**, 381-383 (1978).
7. Brooks, C., Hart, S. R. & Wendt, I. *Rev. Geophys. Space Phys.* **10**, 551-577 (1972).
8. Roddick, J. C. & Compston, W. *Earth planet. Sci. Lett.* **34**, 238-246 (1977).
9. Pankhurst, R. J. & O'Nions, R. K. *Earth planet. Sci. Lett.* **12**, 127-136 (1973).
10. York, D. *Earth planet. Sci. Lett.* **5**, 320-324 (1969).

Table 1 Mean numbers of correct identification as a function of type of characters and of visual field

	Pictograms	Phonograms
Right visual field	5.20	5.60
Left visual field	13.30	14.20

The term 'lateralisation' refers to the specialisation of the left or right hemisphere of the brain for different functions. The rationale behind the visual hemi-field experiment is as follows. When a subject looks at a fixation point in the centre of a lighted square within a tachistoscope, each visual hemi-field projects to the contralateral hemisphere. So, for example, stimuli presented to the RVF are first processed in the left hemisphere. If language is indeed processed in the left hemisphere, then verbal stimuli presented to the RVF should take less time to respond than when the same materials are presented to the LVF. The delay in reaction time is attributed to the time it takes to transfer information from one hemisphere to the other. Or the experimenter can shorten the exposure duration so that the subjects are expected to make identification errors. Depending on the pattern of such an accuracy measure (that is RVF or LVF superiority) and on the materials used, specific functions of the left or right hemispheres can be inferred.

orthography-specific property; rather, it reflects a task-specific property of our cerebral hemispheric processing.

In our first experiment, 80 stimulus characters were selected from established norms of commonly used Chinese characters (both abstract and concrete terms were used). Two types of characters were selected: phonograms, which contain a phonetic component as a clue to the sound of the characters, and pictograms, which are pictographic in origin⁸. Figure 1 presents some examples of these two types of character. The subjects were 20 Chinese students (all right-handed) enrolled in the graduate school of the University of California, Riverside. The stimulus characters were presented one at a time either in the RVF or on the LVF by means of a three-channel tachistoscope. Each character was always preceded by an auditory signal plus a 500-ms fixation point presented in the centre of the screen. The duration of character presentation was adjusted for each subject according to his/her recognition threshold determined before the experimental session (range 20–75 ms; mean 40 ms). The subject was asked to identify verbally the stimulus character as soon as it was presented. The entire experiment forms a 2 (phonogram compared with pictogram) $\times$ 2 (RVF compared with LVF) design with both factors as repeated measures, and the dependent measure was the number of correct identifications. The results are summarised in Table 1. A 2 $\times$ 2 analysis of variance (ANOVA) was performed on the data. Only one main effect of the visual field is statistically significant, $F(1, 19) = 49.56$, $P < 0.001$. Apparently a strong LVF-RH superiority effect was obtained regardless of whether the characters contained any phonetic cue.

Although the result of a RH processing is clearcut, its implication for reading is less clear. Modern Chinese tends to be multiple-syllable, and so the perceptual unit in reading may be much larger than the single characters. Thus, a major task in reading is to generate meaning by putting together several characters to form meaningful terms. The second experiment, therefore, was conducted to investigate whether the results obtained in the previous experiment would be generalised to the tachistoscopic recognition of multiple-character terms. The procedure was the same as for the first experiment except that the stimulus terms were always two characters arranged vertically in either LVF or RVF of the tachistoscope. The results showed that of the 80 stimulus terms presented (40 for each field) the numbers of correct identification were 36.5 and 22.5 for stimulus terms presented in the RVF and LVF, respectively. The difference is statistically significant, $t(19) = 7.49$, $P < 0.001$. This result suggested a LH dominance, a complete reversal of the previous experiment. Therefore, the locus of the cerebral lateralisation cannot be on the logographic symbols *per se*; rather, it must be on the type of task a subject has to perform to meet the experimental requirement.

One may still object to the result of the second experiment on the basis that a LVF-RH effect as observed in the first experiment might be overridden by the LH function (it is conceivable that more effort is required to pronounce two characters) of motor programming in the oral report. To circumvent this objection, a third experiment was run to investigate the visual lateralisation effect with the subject making a semantic decision on multiple-character terms. Twenty Chinese graduate students were asked to make a 'yes/no' response as quickly and as accurately as possible to stimulus terms briefly exposed in either the RVF or LVF of the tachistoscope. They were asked to press the 'yes' key only if the characters formed a meaningful term, otherwise to press the 'no' key. The right-hand or left-hand responding to the 'yes' or 'no' key was counterbalanced across subjects. We also manipulated the imagery value by dividing the stimulus terms into 16 high and 16 low imagery terms. Thirty-two meaningless character strings were constructed to balance the positive and negative instances. The dependent measure was the reaction time (RT) required to make a correct decision. The results are summarised in Table 2.

First, subjects took much longer to make a negative response than a positive response and the data on negative responses were much noisier than those on the positive responses. Because our interest is in the positive response in various experimental conditions, ANOVA for a 2 (RVF compared with LVF) $\times$ 2 (high compared with low imagery) factorial design with both factors as repeated measures was performed on RTs of the correct positive responses. The main effect of imagery is significant with the high imagery terms having a 62-ms advantage over the low imagery terms, $F(1, 19) = 5.69$, $P < 0.05$. This is consistent with previous findings⁹. The main effect of visual field is also highly significant, $F(1, 19) = 14.52$, $P < 0.002$, with RVF performance having a 104.1-ms advantage over the LVF performance. In other words,

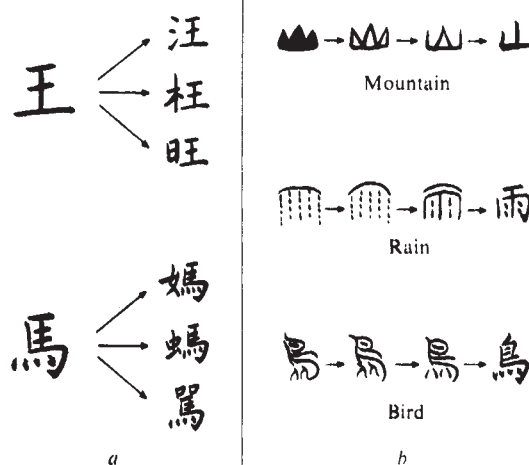


Fig. 1 a, Examples of Chinese phonograms. In the upper panel, the character on the left hand side is the base character and is pronounced as 'wáng' (meaning 'king'). The three characters on the right are derivatives which contain the base character as a clue to their pronunciation. In fact, they are pronounced as 'wǎng', 'wǎng' and 'wàng' from top to bottom, meaning 'the barking sound of dogs' (or alternatively, 'deep and wide'), 'not straight', and 'prosperity' for the three characters, respectively. In the lower panel, the base character on the left is pronounced as 'mǎ'. It means 'horse' and it is a pictogram by itself⁸. Similarly, the three derivative characters on the right are pronounced as 'mā', 'mǎ' and 'mà', meaning 'mother', 'ant' and 'to scold', respectively. Thus, if a reader knows how to pronounce the base characters, he can guess the pronunciation of the derived phonograms which contain the base character as its partial component. However, one should be cautious in making generalisations because in a lot of cases the base character only gives a clue to how a particular phonogram sounds (sometimes the clue refers only to the vowel ending) and the tonal patterns (—, /, \, \) are not included. b, Examples of the pictograms and their transformation through hundreds of years.

Table 2 Mean reaction times for correct decision as a function of imagery value and of visual field

	Positive response		Negative response
	High imagery	Low imagery	
Right visual field	827.82	863.18	1,382.95
Left visual field	906.34	993.52	1,391.32

Responses are given in ms.

the results showed a RVF-LH superiority effect, replicating the result found in the second experiment.

Superficially, this result of a LH dominance seems to be in conflict with those of the first experiment. However, these data are compatible with the current view of cerebral organisation which maintains that, for most right-handed people, the left hemisphere is better at handling sequential and analytical tasks while the right hemisphere is better at handling holistic pattern recognition tasks¹⁰. Because the experimental task in the first experiment emphasises a holistic recognition of the single characters and because the experimental task in the second and third experiments emphasises putting character strings together to get at the lexemes of the terms, the results of these three experiments not only do not conflict with one another, but rather they support the differential-function views of cerebral organisation. Table 2 shows that the data on the imagery variable are consistent with this interpretation.

According to this view, the right hemisphere is superior in perceiving overall form. Clearly, the LVF-RH advantage for single-character recognition must have derived from this source, although final pronunciation of the meaning of the character must be mediated by the left hemisphere. This implies that the left hemisphere's articulatory superiority is outweighed by the right hemisphere's form-perception superiority for the single-character test. It seems evident that the reading and interpretation of Chinese characters calls on the specialised functions of both sides of the brain, and that whether one or the other side manifests superiority in a particular task depends on the balance of cognitive requirement.

Received 25 April; accepted 19 September 1979.

- Hardyck, C., Tzeng, O. J. L. & Wang, W. S.-Y. *Nature* **269**, 705-707 (1977).
- Hirata, K. & Osaka, R. *Psychologia* **10**, 7-18 (1967).
- Hatta, T. *Neuropsychologia* **15**, 685-688 (1977).
- Sasanuma, S., Itoh, M., Mori, K. & Kobayashi, Y. *Neuropsychologia* **15**, 547-553 (1977).
- Sasanuma, S. *Cortex* **10**, 89-97 (1974).
- Wood, C. C., Goff, W. R. & Day, R. S. *Science* **173**, 1248-1251 (1971).
- Tzeng, O. J. L., Hung, D. L. & Wang, W. S.-Y. *J. exp. Psychol. Hum. Learning Memory* **3**, 621-630 (1977).
- Wang, W. S.-Y. *Scientist Am.* **228**, 51-60 (1973).
- Pavio, A. *Imagery and Verbal Processes* (Holt, Rinehart and Winston, New York, 1971).
- Patterson, K. & Bradshaw, J. L. *J. exp. Psychol. Hum. Perception Performance* **1**, 246-252 (1975).

Spider feeding behaviour optimises dietary essential amino acid composition

Matthew H. Greenstone*

Department of Zoology, University of California, Berkeley, California 94720

Classical models of foraging behaviour¹⁻³ assume that feeding patterns maximising the net caloric gain per unit time or minimising the time spent in foraging will be favoured by natural selection. There is some evidence for this behaviour in some foragers⁴, mostly higher vertebrates for which time and energy constraints are frequently paramount. However, while these models may suffice under some circumstances, nutritional

information from the veterinary, wildlife management, and entomological literature shows that an animal requires a diet containing all necessary nutrients to be truly fit (physiologically and evolutionarily). Therefore unless animals are very severely constrained energetically, foraging behaviour should lead to nutritional as well as caloric optimisation. Here I show that free-living wolf spiders will tend to prey on three species in proportions which optimise the proportions of the essential amino acids they provide in the diet.

The small (<50 mg) lycosid, *Pardosa ramulosa* (McCook), is associated in California with standing water, where it feeds predominantly on aquatic insects⁵. In the Petaluma Marsh in Sonoma County the spiders are found at isolated brackish pools. Eighty per cent of their diet comprises three insect species, the shore fly, *Ephydra riparia*, the waterboatman, *Trichocorixa reticulata*, and the mosquito, *Aedes dorsalis*⁵.

Information on short-term dietary selectivity in the presence of a particular set of prey availabilities was obtained by statistical analysis of qualitative data on spider stomach contents, which were identified with a passive haemagglutination inhibition assay⁶. Because the assay has perfect specificity, each spider may be unambiguously scored either positive or negative for the presence of each of the three insect species. On this basis, each spider in a sample may be assigned to one of the eight cells of a 2 × 2 × 2 contingency table for a test of independence among the prey species in the diet, using the marginal totals to compute the expected numbers of spiders in each cell. The null hypothesis

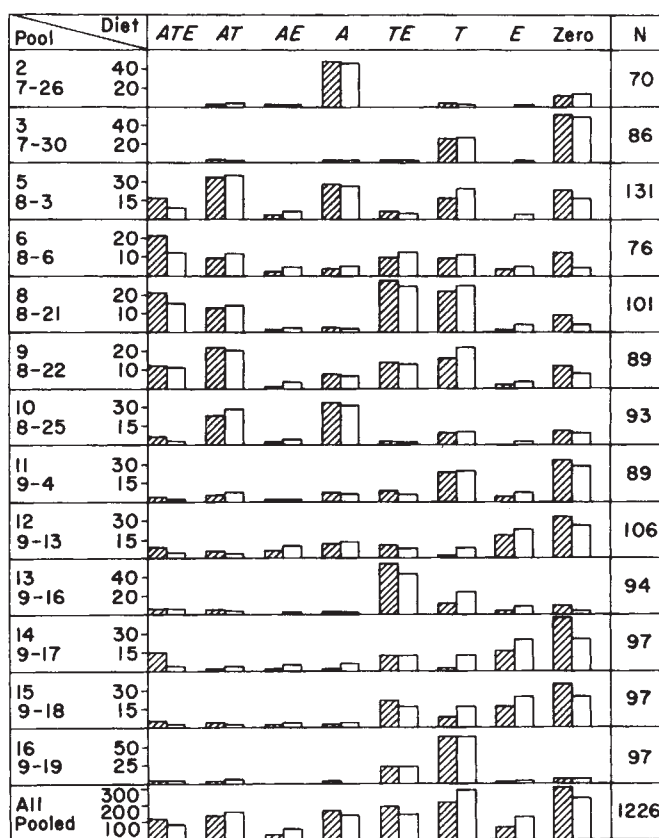


Fig. 1 Observed (hatched bars) and expected (open bars) numbers of spiders with each of the eight possible kinds of diets. Symbols are as in Table 1. A certain proportion of the spiders in the contingency table (the 'zero' category) are expected to be negative for all three species. In addition to these animals, a subset of the sample, whose identities are unknown, will not be feeding because they are preparing to moult¹³. Three per cent of the spiders in a laboratory population were not eating at any given time because of preparation for moult (unpublished data). Therefore, 3% of each sample, as zeroes, were omitted from the analysis in order to avoid an artificial inflation of the observed number of zeroes in the contingency table.

* Present address: Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92717.

Table 1 *G* values and their associated probabilities for the contingency table tests of independence of the prey species in the diet

Pool	T×A Independence	T×E Independence	E×A Independence	A×T×E Interaction	A×T×E Independence	N
2	1.499	0.118	0.523	0.000	2.141	70
3	0.139	1.997	0.048	0.000	2.184	86
5	1.367	10.906§	1.912	4.463†	18.647§	131
6	6.802†	6.802†	3.375	0.032	17.012§	76
8	1.107	10.684§	0.892	0.371	13.055§	101
9	1.553	4.637†	0.451	0.045	6.686	89
10	0.031	3.594	0.305	1.352	5.283	93
11	0.009	3.974†	1.408	0.007	5.399	89
12	4.070†	8.412§	0.006	1.312	13.799§	106
13	0.370	19.111§	0.697	0.150	20.327§	94
14	24.821§	27.143§	18.481§	0.001	70.445§	97
15	3.607	12.946§	0.537	0.767	17.857§	97
16	1.394	0.163	0.163	0.045	0.885	97
All samples pooled:	1.461	119.992§	0.123	1.105	122.574§	1226

A = *Aedes*, T = *Trichocorixa*, and E = *Ephydra*; N = sample size.

† $P < 0.05$, ‡ $P < 0.01$, § $P < 0.001$.

can be rejected in either of two ways. Positive association is consistent with a hypothesis of nutrient optimisation in which certain proportions of several species provide more nearly optimal proportions of nutrients than does any single species, or in which taking a variety of species minimises the consumption of specific noxious defensive compounds from each one⁷. Negative association of prey species in the diet would suggest specialisation on a particular species, because it provides the greatest net caloric gain or a more nearly optimal nutrient balance than any combination of species. It must be emphasised that the availabilities and dispersions of prey species on any given date and any prey species differences in catchability, visibility, and encounter rate are completely irrelevant to this analysis since they are subsumed in the marginal totals of the contingency tables.

Thirteen samples of about a hundred spiders each were collected during the summer of 1974, and the observed and expected values for the contingency tables derived from the immunological analysis of their stomach contents were compared by means of the *G* test⁸ (Table 1). The pooled data, and nine of the thirteen individual samples, show significant departure from independence; in the majority of samples this is due entirely to *Trichocorixa* × *Ephydra* heterogeneity. The observed and expected values for the eight cells of each table are shown in Figure 1. All of the significant pools show an excess of triple positives and, with the single exception of pool 6, an excess of spiders positive for *Trichocorixa* and *Ephydra* and negative for *Aedes*, and deficiencies of *Trichocorixa* and *Ephydra* singles. These data indicate dietary mixing.

Because of the importance of essential amino acids in the diets of arthropods^{9,10}, I decided to test the hypothesis that the

significant positive associations of prey species reflect attempts by the spiders to provide the most nearly optimal proportions of essential amino acids in the diet. The null hypothesis is that the proportions of amino acids in the diets of the average spiders in the four nonsignificant samples (pools 2, 3, 10 and 16) are not significantly more or less similar to the optimal proportions than are the proportions of amino acids in the diets of average spiders in the nine significant samples. The test hypothesis is that the proportions of essential amino acids in the diets of the average spiders of the nine significant samples are significantly more similar to the optimal proportions than are the proportions of essential amino acids in the diets of the average spiders of the four non-significant samples.

The optimal proportions of essential amino acids are assumed to be those in the spiders themselves. (This is obviously a first approximation since rates of utilisation may not be the same for all amino acids.) The proportion of the *i*th amino acid in the diet of the average spider in a sample is given by

$$p_i = \frac{\sum_j \lambda_j \mu_{ij}}{\sum_i \sum_j \lambda_j \mu_{ij}} \quad (1)$$

where λ_j is the mean number of individuals of prey species *j* per spider stomach in the sample and μ_{ij} is the weight of amino acid *i* in an individual of prey species *j*. The proportions of the 10 essential amino acids in the spiders and their prey (Table 2) were determined by dansylation and chromatography¹¹ of appropriate extracts. The λ_j values were estimated from the immunological data by assuming that the number of individuals of each prey species per spider stomach is a Poisson variate¹². Then the probability that a stomach contains *r* items of species *j* is given by

$$p_{rj} = \frac{\lambda_j^r e^{-\lambda_j}}{r!}$$

from which it follows that

$$\lambda_j = -\ln(p_{0j})$$

where p_{0j} is the proportion of spiders in the sample which is negative for species *j*, after correction for animals not feeding because of preparation for moult¹³. The resulting λ_j values were corrected for differences in the length of time during which individuals of the three species remain detectable in the stomach^{5,14}. These corrected λ_j values (Table 3) were entered into equation (1) to compute the proportion of each of the ten essential amino acids in the diet of the average spider in each sample.

The similarity between each of the 13 average diets and the optimum diet was indexed using vector dot products. Since the proportions in any diet or in the optimum sum to 1.0, the sets of proportions can be represented as probability vectors. The dot

Table 2 Proportions of the essential amino acids in the spiders and their prey

Amino acid	<i>Pardosa</i>	<i>Aedes</i>	<i>Trichocorixa</i>	<i>Ephydra</i>
Arg	0.0280	0.0437	0.0514	0.0880
His	0.0093	0.0299	0.0363	0.0281
Ile	0.1701	0.1764	0.1441	0.2168
Leu	0.2295	0.2556	0.1425	0.1735
Lys	0.0560	0.0507	0.1542	0.0545
Met	0.0147	0.0151	0.0346	0.0352
Phe	0.1054	0.2046	0.0832	0.1022
Thr	0.0869	0.0398	0.0793	0.0683
Val	0.1889	0.1827	0.1492	0.1590
Trp	0.1113	0.0014	0.1251	0.0744

Spider extracts were prepared by grinding whole animals in distilled water. Haemolymph alone was used for the insect comparisons (see ref. 6 for method of extraction) because the spiders do not consume their prey intact and therefore do not ingest skeletal material.

product of two vectors, α and β , is¹⁵

$$\alpha \cdot \beta = |\alpha||\beta| \cos \theta$$

In the case of probability vectors,

$$\theta = \arccos \frac{(\sum p_{di} p_{oi})}{(\sum p_{di}^2 \sum p_{oi}^2)^{1/2}} \quad (2)$$

where p_{di} is the proportion of the i th amino acid in the diet which is being compared to the optimum and p_{oi} is the proportion of the i th amino acid in the optimum diet. When two vectors are identical $\theta = 0^\circ$. The more dissimilar they are the larger θ becomes, up to a maximum of 90° (ref. 16).

The θ values computed for each pool's average diet versus the optimum using equation (2) are presented in Table 3. There is no overlap between the θ values of the two sets of comparisons, all of the non-significant pools' diets' comparisons showing θ values between 19.24° and 21.92° , and all of the significant pools' diets' comparisons showing θ values between 13.11° and 18.18° , indicating diets with a closer similarity to the optimum. This difference is highly significant by the Mann-Whitney U -test¹⁷, ($U = 0$, $n_1 = 4$, $n_2 = 9$, $P < 0.001$), and strongly supports the nutritional hypothesis.

Dietary mixing has been reported in rats¹⁸, deer mice¹⁹, and flour beetles²⁰ in the laboratory, and in songbirds in the field²¹, but has not yet been described in free-living invertebrate predators. The fact that a bug (*Trichocorixa*) and a fly (*Ephydra*) were most often associated in the spiders' diet agrees with field observations that the insects most heavily represented in the diets of European^{22,23} and North American^{5,24} *Pardosa* species belong to the orders Homoptera, Hemiptera and Diptera. Perhaps these spiders have adopted this particular form of polyphagy for nutritional reasons. Such behaviour probably has selective value, since, in the laboratory at least, survival^{25,26} and fecundity²⁶ are higher in *Pardosa* species raised on multiple species than on single species diets.

Recognition of the importance of nutrients besides calories has appeared in the optimal foraging literature in the past few years²⁷⁻³². It has been suggested, for example, that protein nitrogen^{27,28} or other nutrients²⁹ be substituted for calories as the currency for optimisation. In reality the decisions made by foraging animals will probably be better represented by more complex models entailing compromises in the simultaneous optimisation of caloric and other nutritional factors and, as shown in the present study, concerned with both the quality and quantity of nutrients.

I thank Robert Colwell for useful discussions, Irene Baker for performing amino acid analyses, Steve Selvin, Robert Sokal and Howard Tucker for statistical advice and Paul Summarco for

performing G -tests. I also thank Charles Mitter and Robert Sokal for helpful comments. This work was supported in part by NSF grant GB 41178 and a Biomedical Research Support grant from the School of Biological Sciences, University of California, Irvine.

Received 10 July; accepted 14 September 1979.

1. Emlen, J. M. *Am. Nat.* **100**, 611-617 (1966).
2. MacArthur, R. H. & Pianka, E. R. *Am. Nat.* **100**, 603-609 (1966).
3. Schoener, T. W. *A. Rev. ecol. Syst.* **2**, 369-404 (1971).
4. Pyke, G. H., Pulliam, H. R. & Charnov, E. L. *Q. Rev. Biol.* **52**, 137-154 (1977).
5. Greenstone, M. H. thesis, Univ. Calif. Berkeley (1976).
6. Greenstone, M. H. *J. appl. Ecol.* **14**, 457-464 (1977).
7. Freeland, W. J. & Janzen, D. H. *Am. Nat.* **108**, 269-289 (1974).
8. Sokal, R. R. & Rohlf, F. J. *Biometry*, 589 (Freeman, San Francisco, 1969).
9. House, H. L. *A. Rev. Biochem.* **31**, 653-672 (1962).
10. Dadd, R. H. *A. Rev. Ent.* **18**, 381-420 (1973).
11. Baker, I. & Baker, H. G. *New Phytol.* **76**, 87-98 (1976).
12. Nakamura, M. & Nakamura, K. *Oecologia* **27**, 97-116 (1977).
13. Hardman, J. M. thesis, Simon Fraser Univ. (1972).
14. Custer, T. W. & Pitelka, F. A. *Condor* **77**, 210-212 (1975).
15. Campbell, H. G. *Matrices with Applications*, 14 (Appleton-Century Crofts, New York, 1968).
16. Fuentes, E. R. *Ecology* **57**, 3-17 (1976).
17. Siegel, S. *Non-parametric Statistics for the Behavioral Sciences*, 116-127 (McGraw-Hill, New York, 1956).
18. Barnett, S. A. *The Rat* 61 (University of Chicago Press, 1975).
19. Holling, C. S. *Mem. ent. Soc. Canada* **45**, 1-60 (1965).
20. Waldbauer, G. P. & Bhattacharya, A. K. *J. Insect. Physiol.* **19**, 407-418 (1973).
21. Tinbergen, L. *Archs neerl. Zool.* **13**, 266-379 (1960).
22. Edgar, W. D. *J. Zool. Lond.* **159**, 405-411 (1969); *Neth J. Zool.* **20**, 487-491 (1970).
23. Hallander, H. *Oikos* **21**, 337-340 (1970).
24. Yeagan, K. V. *Env. Ent.* **4**, 137-141 (1975).
25. Miyashita, K. *Appl. Ent. Zool.* **3**, 81-88 (1968).
26. Hydhorn, S. B. thesis, Univ. Calif. Berkeley (1977).
27. Slansky, F., Jr., & Feeney, P. *Ecol. Monogr.* **47**, 209-228 (1977).
28. White, T. C. R. *Oecologia* **33**, 71-86 (1978).
29. Stenseth, N. C. & Hansson, L. *Am. Nat.* **113**, 373-389 (1979).
30. Marten, G. G. *Ecology* **54**, 92-101 (1973).
31. Pulliam, H. R. *Am. Nat.* **109**, 765-768 (1975).
32. Westoby, M. *Am. Nat.* **108**, 290-304 (1974).

Chronic intracerebroventricular infusion of insulin reduces food intake and body weight of baboons

Stephen C. Woods, Elizabeth C. Lotter, L. David McKay & Daniel Porte Jr

Departments of Psychology and of Medicine and the Regional Primate Center, University of Washington, and the Seattle Veterans Administration Hospital, Seattle, Washington 98195

Body adiposity is normally maintained within rigid limits¹⁻³. Although it is not clear that this regulation fits a strict negative feedback pattern, animals do maintain a relatively constant body adiposity⁴. It has been postulated that this regulation is mediated by some signal which informs centres controlling food intake, probably located in the brain, as to the present state of adiposity^{5,6}. The identity of the signal is unknown, but the direct correlation between body adiposity and basal insulin levels in the plasma⁷⁻⁹, suggests insulin as a possible candidate. This hormone is present in the cerebrospinal fluid (CSF) of many species¹⁰⁻¹³, and is a slow integral over time of the level within the plasma¹⁴. Thus, the level of insulin in the CSF is relatively resistant to short-term plasma fluctuations of insulin. Obese humans have higher levels of CSF insulin than lean controls and the CSF insulin level of both obese and lean humans is reduced proportionately after a prolonged fast¹⁵. We have therefore postulated¹⁶ that the feedback system responding to body adiposity uses the concentration of insulin in the CSF as a major signal. Additional support for such a role is found in recent reports that insulin receptors are present in several regions of the brain and spinal cord¹⁷⁻²⁰. We now present additional evidence for our hypothesis by showing that in baboons the infusion of exogenous insulin into the CSF elicits a reliable and predictable decrease in food intake and body weight.

Table 3 λ_i Values (mean numbers of individuals of each species eaten per spider per unit time) and θ values from vector dot products, for all samples

Pool	λ_A	λ_T	λ_E	θ
2	1.50	0.12	0.01	21.97°
3	0.02	1.02	0.01	18.57°
5	1.02	1.56	0.23	18.18°
6	0.72	2.04	0.72	15.49°
8	0.51	3.50	0.67	15.41°
9	0.68	2.46	0.39	15.22°
10	1.41	1.18	0.12	19.84°
11	0.25	1.26	0.23	14.18°
12	0.42	0.62	0.60	16.54°
13	0.13	3.48	1.02	14.62°
14	0.23	0.84	0.62	14.12°
15	0.14	1.02	0.63	13.11°
16	0.03	4.24	0.31	19.24°

A = *Aedes*, T = *Trichocorixa*, E = *Ephydra*. λ_i values were corrected for differences in detectability between the three species^{5,14}. Detectabilities were determined in the laboratory with controlled feedings under field temperature and light conditions⁵.

To test this hypothesis we have developed a baboon model in which chronic feeding can be monitored in unanaesthetised animals and in which experimental access to the blood and CSF is maintained. Animals were maintained in standard primate restraining chairs for 3–6-week studies during which solutions of sterile buffered salt solution, 'synthetic CSF' (ref. 21), were infused into the animals' lateral cerebral ventricles with or without insulin or glucagon. To do this, an acrylic pod with six cannula guide sleeves was stereotactically fixed to the skull while the animal was anaesthetised with halothane²². The guide sleeves were cut so that their tips would lie at the dorsal surface of the lateral cerebral ventricle, and X rays were used to compare their positions with the plates in Riche²³.

All animals were adapted to a commercially available liquid diet (Ensure, Ross, containing per l approximately 37 g fat, 37 g protein and 145 g carbohydrate with additional vitamins and minerals) which was freely available in stainless steel drinking cans placed on the chairs from 8.30 a.m. to 3.30 p.m. each day. Water only was available from 3.30 p.m. to 8.30 a.m. Food intake was determined by periodic volumetric determinations of remaining diet during the feeding interval, with addition of supplemental diet as appropriate.

To verify that infusions through the cannula actually reached the CSF, we routinely injected each animal with 5 μ g of angiotensin II through a cannula 1–2 mm longer than the guide sleeve and determined subsequent water intake. An increase in water intake of more than 50 ml over the 30 min following the injection relative to the water intake on a control day was taken as indicative that the cannula had access to the CSF. The mean water intake over the 30-min interval immediately following the injection of the angiotensin II was 277 ml greater than that on a control day when only the vehicle was injected.

The intraventricular cannulas were infused at a constant rate between 1 and 1.5 ml per day. Infusions of synthetic CSF had no apparent effect on food intake, body weight, basal plasma insulin or glucose levels for prolonged intervals (see Table 1 and Fig. 1). After a control or baseline interval, insulin (bovine/porcine insulin, Lilly) was added to the infusate at doses of 1, 10 or 100 μ U per kg per day. A given baboon received each

dose of insulin only once, and the order of the various conditions was random. The lowest dose of insulin was chosen because we reasoned that if this was given in a single dose instead of being infused over a 24-h period it would approximately double the concentration of insulin in the CSF. This assumed that the level of insulin within the CSF of the baboon is approximately 25% of the level in the plasma, as it is in other species^{10–14}, and that an average baboon weighs 15 kg and has a ventricular volume around 1 ml. The increment of CSF insulin achieved with the lowest infusion dose was probably negligible as it was actually infused over a 24-h interval and the half life of insulin in the CSF of the dog is around about 1.5 h (ref. 12).

Results of this procedure are summarised in Table 1, which contains the mean values for a 2-week baseline interval in which only synthetic CSF was infused and analogous mean values for the intervals after with the infusion of insulin or glucagon (Lilly). Figure 1 depicts the daily food intake for these conditions. Although there was considerable daily variation in food intake, there was a significant, dose-related decrease in food intake when insulin was added to the infusate. The data were analysed using non-parametric statistics because there was often heterogeneity of variance between conditions and because not every baboon was used in every condition. For the analyses, a mean value was obtained for each animal over the 14-day baseline period. A second overall mean value was obtained for the interval when the hormone was actually infused. These data were then analysed using an overall Kruskal–Wallis test with individual conditions compared by the Mann–Whitney test²⁴; $P < 0.05$ indicated significance. It should be clear that the analysis was very conservative, as the use of some other parameter (for example, a mean of the last 5 days' values in each condition) would have resulted in a larger change, thus giving a greater probability to achieve a significant effect.

There were no reliable differences between baseline parameters for the four conditions (three insulin doses and one glucagon dose). For food intake (Table 1) there was a significant reduction during infusion of both the two larger doses of insulin, and the intake during the largest dose (100 μ U per kg per day) was significantly less than that during the intermediate dose

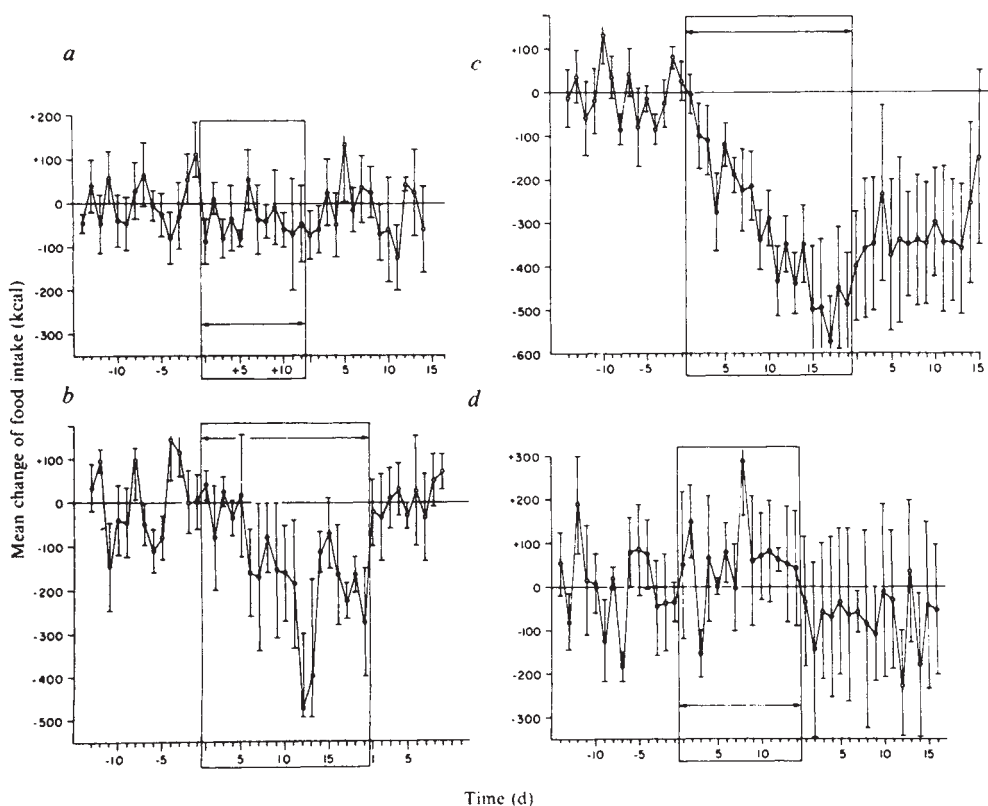


Fig. 1 Daily food intake values for the four experiments. Points and bars represent the mean $\pm$ 1 s.e.m. Open symbols represent control intervals when 'synthetic CSF' was infused; solid symbols represent intervals when insulin at one of three doses (a, 1 μ U per kg per day; b, 10 μ U per kg per day; c, 100 μ U per kg per day) or glucagon (d, 2.35 ng per kg per day) was infused. In a the mean baseline intake was 821 kcal d⁻¹ (n=5), in b 860 kcal d⁻¹ (n=4), in c 781 kcal d⁻¹ (n=6) and in d 726 kcal d⁻¹ (n=3). Times of hormone infusions are boxed.

Table 1 Food intake, plasma insulin, plasma glucose and body weight for the baseline and hormone-infusion portions of the experiment

	Mean food intake (kcal)*	Mean plasma insulin ($\mu\text{U ml}^{-1}$)*	Mean plasma glucose (mg ml^{-1})*	Mean body weight (kg)*	Final body weight (kg)†
Baseline	821 $\pm$ 121	71.5 $\pm$ 11.1	86.3 $\pm$ 1.9	17.8 $\pm$ 2.6	
Insulin (1 μU per kg per day)	786 $\pm$ 119	73.8 $\pm$ 11.4	85.3 $\pm$ 1.3	17.8 $\pm$ 2.7	17.8 $\pm$ 3.0
Change (n = 5)	-35 $\pm$ 12	+2.3 $\pm$ 5.5	-1.0 $\pm$ 2.4	+0.04 $\pm$ 0.08	-0.01 $\pm$ 0.08
Baseline	860 $\pm$ 135	65.1 $\pm$ 11.6	84.3 $\pm$ 1.2	18.8 $\pm$ 3.3	
Insulin (10 μU per kg per day)	702 $\pm$ 97	66.1 $\pm$ 8.4	86.0 $\pm$ 2.7	18.5 $\pm$ 3.2	18.4 $\pm$ 3.0
Change (n = 4)	-158 $\pm$ 57‡	+1.0 $\pm$ 6.0	+1.7 $\pm$ 3.9	-0.28 $\pm$ 0.15	-0.37 $\pm$ 0.12‡
Baseline	781 $\pm$ 99	67.2 $\pm$ 3.9	81.0 $\pm$ 4.0	17.8 $\pm$ 2.3	
Insulin (100 μU per kg per day)	462 $\pm$ 99	63.5 $\pm$ 4.4	82.5 $\pm$ 1.5	17.4 $\pm$ 2.2	17.0 $\pm$ 2.4
Change (n = 6)	-319 $\pm$ 45§	-3.9 $\pm$ 0.5	+1.5 $\pm$ 2.5	-0.33 $\pm$ 0.12‡	-0.81 $\pm$ 0.09‡
Baseline	726 $\pm$ 104	46.3 $\pm$ 2.1	87.5 $\pm$ 4.5	14.5 $\pm$ 1.2	
Glucagon	788 $\pm$ 60	52.9 $\pm$ 10.3	87.5 $\pm$ 2.5	14.6 $\pm$ 1.3	14.7 $\pm$ 1.6
Change (n = 3)	+62 $\pm$ 61	+6.6 $\pm$ 8.3	0 $\pm$ 2.0	+0.10 $\pm$ 0.06	+0.19 $\pm$ 0.04

* These values represent an average (mean $\pm$ s.e.m.) for the entire infusion interval.

† These values are those of the final day of hormone infusion and are indicative of the total weight change achieved.

‡ Significantly different than the analogous value during the lowest dose of insulin.

§ Significantly different than the analogous value during the intermediate dose of insulin.

|| Significantly different than the analogous value during the highest dose of insulin.

(10 μU per kg per day). The food intake during the glucagon infusion did not differ significantly from that during the baseline period, nor from that during the infusion of the smallest dose of insulin. It was significantly greater than the intake during the infusion of the largest dose of insulin. With respect to body weight, only the largest infusion of insulin resulted in a significant decrease relative to the baseline or to the smallest dose. The body weight during the glucagon infusion was not consistently different from that during the baseline but significantly larger than that obtained during the infusion of the largest dose of insulin. There were no differences in plasma insulin (as determined by immunoassay)²⁵ or plasma glucose (as determined by the glucose oxidase method) during any hormone infusion interval relative to control. There was a significant dose-related reduction in body weight, but no reliable change in plasma insulin or glucose during these intervals (Table 1), suggesting that leakage of infused insulin from the CSF to the blood was minimal.

The weight losses at the end of the sessions during which the two larger insulin doses were infused (see last column in Table 1) were approximately what one would expect based on the decrease in food intake. If one assumes that 1 g of stored fat yields about 7 kcal (ref. 1), the mean daily decrease of 158 kcal for 19 days (based on the figures for the 10 μU per kg per day dose) would be expected to cause a mean decrease of 429 g of body weight. The actual decrease was 374 g. Similarly, the expected decrease in weight in animals given the largest insulin dose for 19 days would be 866 g, whereas the actual weight loss was 812 g.

At the end of an experimental infusion session in which food intake was depressed, food intake rapidly returned towards the baseline control level. However, the rebound was not complete following the infusion of the largest dose of insulin even after 2 weeks (see Fig. 1).

Glucagon was infused as a peptide control at the same molar dose as the highest effective dose of insulin. As shown in Table 1 and Fig. 1, glucagon infusion at this dose had no consistent effect on food intake, body weight, plasma insulin or plasma glucose. If anything, there was a slight tendency for the baboons to increase their food intake. This was most apparent as a rebound decrease in food intake after the end of the glucagon infusion (Fig. 1), although this result was not statistically significant. However, it is clear that the suppression of food intake observed when equimolar amounts of insulin were infused did not occur with glucagon. This control therefore makes some nonspecific illness-related explanation unlikely, especially when one considers that

glucagon, when given peripherally, reportedly makes animals ill and causes them to eat less^{26,27}.

We have interpreted these results to support the theory¹⁶ that CSF insulin levels may serve as a signal to inform the brain of the present state of adiposity. It may well be that other, analogous systems exist in which CSF-borne peptides influence behaviour. For body weight regulation, the observed system closes the feedback loop between the brain and the endocrine pancreas^{28,29}. Independently of the theoretical interpretation, there are potential practical implications of these findings. We have found that we can reduce the food intake and body weight of baboons by altering the amount of a hormone already present in the CSF. It will be important to determine if this system is operative in man and if there are possible therapeutic implications for human obesity.

This work was supported by NIH grant nos AM 17844, AM 12829 and AM 17047 (University of Washington Diabetes Research Center) and by Veterans Administration Medical Research Funds and the Northwest Regional Primate Center (RR-00166). We thank Keith Vogel, Marilyn Evans, Barbara O'Neill, Howard Beiter, Karen E. Moe, Charles C. Gale and Thomas O. Nelson for their assistance.

Received 25 June; accepted 25 September 1979.

- Bray, G.A. *The Obese Patient* (Saunders, Philadelphia, 1976).
- Woods, S. C., Decker, E. & Vasselli, J. R. *Psychol. Rev.* **81**, 26-43 (1974).
- Assimakopoulos-Jeanet, F. & Jeanrenaud, B. *Clin. Endocr. Metab.* **5**, 337-372 (1976).
- Mrosovsky, N. & Powley, T. L. *Behav. Biol.* **20**, 205-223 (1977).
- Kennedy, G. C. *Proc. R. Soc. B* **140**, 578-592 (1953).
- Hoebel, G. B. A. *Rev. Physiol.* **33**, 533-568 (1971).
- Bagdade, J. D., Bierman, E. L. & Porte, D. Jr *J. clin. Invest.* **46**, 1549-1557 (1967).
- Grant, D. B. *Archs Dis. Childh.* **42**, 375-378 (1967).
- Stephan, F. et al. *Diabetologia* **8**, 196-201 (1972).
- Margolis, R. U. & Altszuler, N. *Nature* **215**, 1375 (1967).
- Anderson, D. K. & Hazelwood, R. C. *J. Physiol., Lond.* **202**, 83-95 (1969).
- Woods, S. C. & Porte, D. Jr *Diabetes* **24**, 905-909 (1975).
- Hill, D. E. et al. *Clin. Res.* **26**, 71A (1978).
- Woods, S. C. & Porte, D. Jr *Am. J. Physiol.* **233**, E331-E334 (1977).
- Owen, O. E. et al. *Diabetes* **19**, 397 (1970); *Metabolism* **23**, 7 (1974).
- Woods, S. C. & Porte, D. Jr in *Hunger: Basic Mechanisms and Clinical Implications* (eds Novin, D. et al.) 273-280 (Raven, New York, 1976); *Adv. metab. Dis.* **9**, 283-312 (1978).
- Havrankova, J., Brownstein, M. & Roth, J. *Nature* **272**, 827-829 (1978).
- Havrankova, J., Schmechel, D., Roth, J. & Brownstein, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5737-5741 (1978).
- Pacold, S. T., Small, E. & Blackard, W. C. *Clin. Res.* **26**, 34A (1978).
- Van Houten, M. et al. *Clin. Res.* **26**, 533A (1978).
- Myers, R. D. *Methods in Psychobiology* Vol. 1 (Academic, New York, 1971).
- Woods, S. C. et al. *J. Neurosci. Meth.* (in the press).
- Riche, D. *Atlas Stereotaxique du Cerveau de Babouin* (Centre National de la Recherche Scientifique, Paris, 1968).
- Hays, W. L. *Statistics for Psychologists*, 637 (Holt, Rinehart and Winston, New York, 1963).
- Porte, D. Jr et al. *J. clin. Invest.* **45**, 228-236 (1966).
- Balagura, S. J. *comp. Physiol. Psychol.* **65**, 30-32 (1968).
- Martin, J. R. & Novin, D. *Physiol. Behav.* **19**, 461-467 (1977).
- Woods, S. C. & Porte, D. Jr *Physiol. Rev.* **54**, 596-619 (1974).
- Porte, D. Jr et al. *Pharmac. Biochem. Behav.* **3**, 127-133 (1975).

Nerve sheaths and motoneurone collateral sprouting

J. R. Slack*, W. G. Hopkins & M. N. Williams

Department of Physiology, School of Medicine, University of Auckland, Private Bag, Auckland, New Zealand

When disease or injury causes partial loss of innervation from a muscle, the remaining axons sprout and form new connections to the denervated muscle fibres¹. Sprouting can occur in two ways: from axon terminals (terminal sprouting) or from the intramuscular axons themselves, probably from the nodes of Ranvier (collateral sprouting)^{1,2}. Terminal sprouting has been induced experimentally using various methods, including partial denervation², nerve conduction block³ and nerve transmission block^{4,5}. A common factor in the induction of terminal sprouting seems to be changes in the surface membrane of muscle fibres^{6,7}; these changes and terminal sprouting are prevented by direct stimulation of the muscle⁵. Collateral sprouting has been induced only by partial denervation and is not prevented by direct stimulation⁵. This has been taken as evidence⁶ for an earlier suggestion² that products of nerve or axon degeneration may be a direct stimulus for collateral sprouting. We report here that axon degeneration products alone are probably not the stimulus for collateral sprouting.

Collateral sprouting in response to partial denervation was examined in a mouse hind limb muscle, tensor fascia latae (TFL), which is normally innervated by 11 axons from lumbar spinal nerves L4 and L5. Mice (15–20 g body weight) were anaesthetised intraperitoneally with chloral hydrate (0.1 ml of 5% solution per 10 g body weight), and partial denervation of TFL was achieved by cutting and redirecting the proximal stump of L4 into the peritoneal cavity. This operation results in the degeneration of a variable number of TFL axons, but usually nine. Muscles were exercised up to 30 days after the operation,

stained by the combined silver and cholinesterase method of Namba *et al.*⁸, mounted whole between two coverslips and examined in the light microscope under a $\times 100$ oil immersion objective. By this method endplates, axons, axon sprouts and perineurial sheaths were visualised throughout the entire muscle. The lengths of sheaths, axons and sprouts were measured using a calibrated 10×10 square graticule in the microscope eyepiece. Sprouts are clearly distinguishable from the pre-existing terminal axons by two criteria. First, the diameter of sprouts and pre-existing axons are significantly different ($0.47 \pm 0.10 \mu\text{m}$, $n = 38$ and $1.79 \pm 0.39 \mu\text{m}$, $n = 95$, mean $\pm$ s.d., respectively). Second, sprouts have uniformly smooth profiles while pre-existing terminal axons possess characteristically gnarled profiles (Fig. 1).

The nerve of TFL branches within the muscle to form smaller bundles of axons (intramuscular nerves) and finally single terminal axons which contact individual muscle fibres. There are no sprouts in muscles removed 1 day after partial denervation. Extensive axonal degeneration is visible within the perineurial sheaths of the intramuscular nerves, some of which contain surviving intact axons. Denervated terminal axon sheaths arise from intramuscular nerve branches and terminate on cholinesterase-stained endplates. Muscles removed from animals two or more days after partial denervation show fine collateral sprouts, which reach denervated muscle fibres by growing down denervated sheaths. Not all the sprouts observed reached cholinesterase-stained endplates. Some (21%, $n = 190$) end in bulbous enlargements similar in appearance to growth cones described by Ramon Y Cajal⁹ in silver-impregnated developing axons. These sprouts which end within perineurial sheaths are most common in muscles partially denervated between two and five days before excision.

Sprouts arise from intact axons close to points at which denervated sheaths branch from partially denervated intramuscular nerves (Fig. 2). In four muscles partially denervated for less than 9 days sprout sites are $\sim 22 \mu\text{m}$ from the branch points ($22.4 \pm 1.8 \mu\text{m}$, $n = 77$). In two muscles partly denervated for 16 and 30 days, sprouts arise $\sim 38 \mu\text{m}$ from the branch points ($38.1 \pm 43.9 \mu\text{m}$, $n = 62$). Sprouts are absent from quite long segments of intramuscular nerve (400–600 μm). If axon degeneration products were the sprouting stimulus, sprouts would be expected throughout these segments, because they

* To whom correspondence should be addressed.

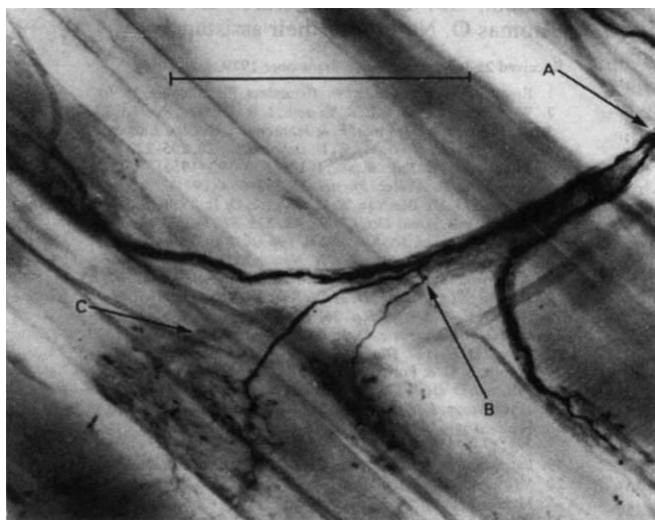


Fig. 1 Light micrograph of combined silver and cholinesterase-stained TFL muscle. The muscle was partially denervated 8 days before excision. Perineurial sheaths enclose two pre-existing terminal axons which arise at arrow A. Cholinesterase enzyme product is visible as speckled deposits over terminal arborisations. A fine collateral sprout arises at arrow B and follows a sheath to a nearby endplate. A denervated sheath (arrow C) arises near B. The axon branch between sprout B and denervated sheath C is one of a small number of profiles (9%, $n = 741$) with a diameter intermediate between that of sprout and pre-existing axon. It may be either a sprout in the process of maturation or an unusually thin pre-existing axon; these are excluded from all sprout statistics. Scale bar, 100 μm .

contain degenerating axons for several days following partial denervation. The degenerating axons are not separated from the intact axons by septa¹⁰ or other diffusion barriers which might prevent access of degeneration products to intact axons (unpublished observations of electron micrographs). The absence of sprouts is not due simply to the absence of potential sprout sites, probably nodes of Ranvier, which occur frequently in these segments (internodal distance = $163 \pm 61 \mu\text{m}$, $n = 31$, from three muscles). Moreover, the sprout-free segments of nerve are

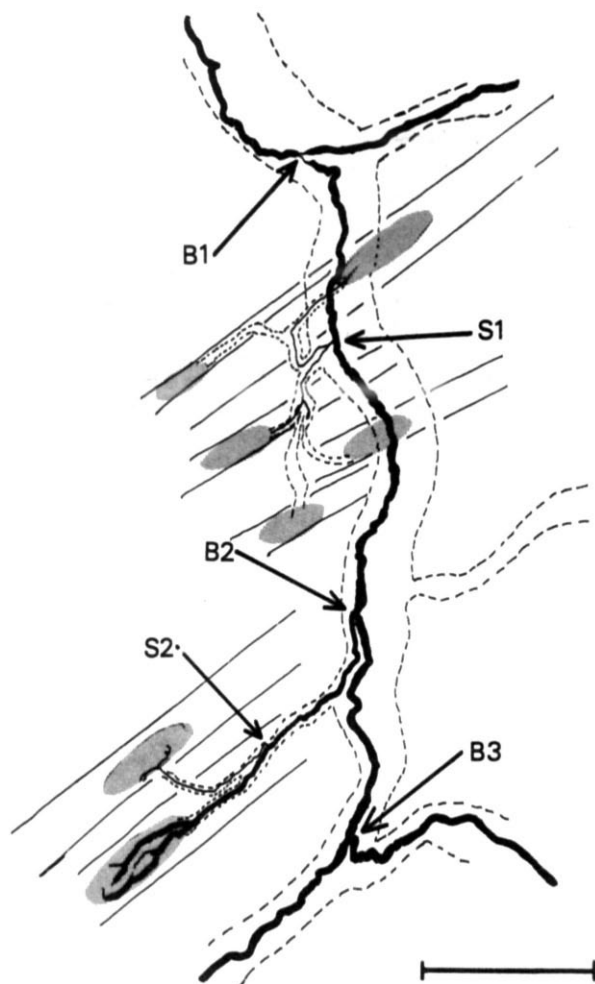


Fig. 2 Camera lucida tracing of a segment of intramuscular nerve from a TFL muscle excised and stained 4 days after partial denervation. Perineurial sheaths enclose degenerating axons (not drawn) and a single intact axon with branch points at B1, B2 and B3. Sprouts arise at S1 and S2 and innervate endplates (shaded areas). A terminal axon arising at B2 also innervates an endplate within the field shown. Note that there is no sprouting at the branch points. In other regions of this muscle, the nodes of Ranvier at branch points are also sprout sites. Scale bar, $100 \mu\text{m}$.

flanked by denervated sheaths containing sprouts, making it unlikely that these segments are for some reason refractory to a sprouting stimulus from degenerating axons.

Denervated terminal axon sheaths become sprout innervated in a sequential manner. We have measured the lengths of sprout-containing sheaths from the endplate to the point where the sprout arises in the intramuscular nerve. In some cases where the sprouts ended in a 'growth cone' short of the cholinesterase-stained endplate we measured the entire distance from the point of sprout outgrowth along the sheath to the endplate. We found that the shortest denervated sheaths are the first to become innervated by sprouts, as shown in Fig. 3.

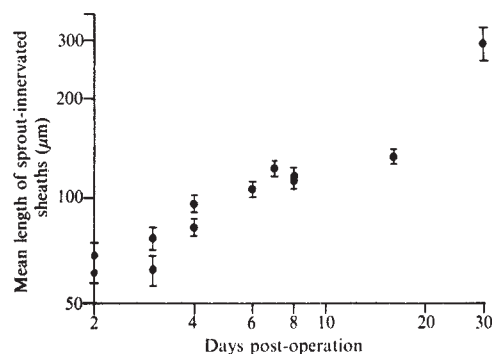


Fig. 3 Mean length of sprout-innervated sheaths plotted as a function of time after partial denervation. Each point represents one muscle and shows the mean $\pm$ s.e.m. for between 21 and 68 sheaths. Mean pre-existing terminal axon length for the 12 muscles in this figure is $130 \pm 8 \mu\text{m}$ (mean $\pm$ s.d.).

These results do not seem to be compatible with a mechanism in which sprouts are initiated simply by the products of nerve or axon degeneration. At least two classes of mechanism for collateral sprout initiation are consistent with our findings. First, the sprouting stimulus might be a diffusible substance in the extracellular space which reaches intact axons either directly across denervated sheaths or via the terminal ends of denervated sheaths. Second, the stimulus might arise from changes in either the denervated Schwann or perineurial sheath cells developing first near the terminal ends of the sheaths, and subsequently spreading back towards the intramuscular nerves containing intact axons.

This work was supported by grants from the MRC of New Zealand and the Auckland Medical Research Foundation.

Received 31 July; accepted 18 September 1979.

1. Edds, MacV. *Q. Rev. Biol.* **28**, 260-276 (1953).
2. Hoffman, H. *Aust. J. exp. Biol. med. Sci.* **28**, 383-397 (1950).
3. Brown, M. C. & Ironton, R. *Nature* **265**, 459-461 (1977).
4. Duchon, C. W. & Strich, S. J. *Q. J. exp. Physiol.* **53**, 84-89 (1968).
5. Ironton, R., Brown, M. C. & Holland R. L. *Brain Res.* **156**, 351-354 (1978).
6. Brown, M. C., Holland, R. L. & Ironton, R. *Nature* **275**, 652-654 (1978).
7. Pestronk, A. & Drachman, D. B. *Science* **199**, 1223-1225 (1978).
8. Namba, T., Nakamura, T. & Grob, D. *Am. J. clin. Path.* **47**, 74-77 (1967).
9. Ramon Y Cajal, S. in *Studies on Vertebrate Neurogenesis*, 46 (Thomas, Illinois, 1960).
10. Burkel, W. E. *Anat. Rec.* **158**, 177-190 (1967).
11. Bradley, W. G. & Thomas, P. K. in *Disorders of Voluntary Muscle* (ed. Walton, J. N.) 239-244 (Churchill-Livingstone, Edinburgh, 1974).

Teratogenic drugs inhibit tumour cell attachment to lectin-coated surfaces

Andrew G. Braun, David J. Emerson
& Bradley B. Nicholson

Harvard Medical School, Department of Radiation Therapy, 50 Binney Street, Boston, Massachusetts 02115

Interactions between embryonic cells are generally thought to have a central role in the control of development¹. When these morphogenic interactions are interrupted by either physical intervention² or genetic defects³, normal development is impaired. In accord with these experiments, specific interactions between embryonic cells have been demonstrated in several *in vitro* systems. Many investigators have described homotypic aggregation of chick embryo cells⁴⁻⁸, and heterotypic specificity has also been described⁹. Because of the importance of morphogenic cell-cell interactions in development it follows

that agents that interfere with these interactions, regardless of the interference mechanism, are potential teratogens. Here we have used a simple *in vitro* cell to surface recognition system in an attempt to screen for potential teratogens. We have found a very high correlation between inhibitory activity in the *in vitro* assay and reported teratogenic activity in human or animal studies. This suggests that many teratogenic agents may act by interfering, in an as yet unknown way, in normal cell to cell interactions.

In this system, tumour cell attachment to concanavalin A-coated plastic surfaces, mediated by a cell-surface carbohydrate-concanavalin A receptor interaction, has been used to mimic some features of the complex interactions to be expected during development. Tumour cells are useful because of their availability in large, homogeneous quantities and ease of preparation. In addition, they have long been known to have properties in common with cells of embryonic origin¹⁰. Of particular relevance is the observation that although normal fibroblasts attach poorly to surfaces derivatised with lectins (ref. 11 and A.G.B., unpublished observations), transformed fibroblasts, tumour cells (refs 11–13 and A.G.B., unpublished observations) and some embryonic cells (P. Spear, J. L. Wang and G. M. Edelman, unpublished, but cited in ref. 25), attach to these surfaces. Both embryo cell aggregation and tumour cell attachment to lectin-coated surfaces require energy^{6,11}. Thus, it is plausible that the attachment of tumour cells to lectin-coated surfaces can serve as a simple *in vitro* model for the morphogenic interactions observed *in vivo*.

Tumour cell attachment to lectin-coated surfaces was quantified by using 1.25-cm diameter polyethylene disks and ascitic mouse ovarian tumour (MOT)¹⁴ labelled with tritiated thymidine *in vivo*. We have found that the tumour cells attach to lectin-derivatised disks rapidly, within 5 min at room temperature. To assay for the inhibitory activity of a drug, tumour cells (1 ml at 5×10^6 per ml) were incubated with the drug, usually for 30 min at 37 °C, poured into a 35-mm polycarbonate Petri dish and several disks submerged into the suspension. As dead cells do not attach to lectin-coated surfaces the drug concentrations used here were kept below overtly toxic levels as determined by trypan exclusion. After a 20 min incubation at room temperature the disks were removed with fine forceps, washed in saline and counted in a scintillation counter. An example of the inhibitory activity of several anaesthetics and tranquilisers is shown in Fig. 1. Similar concentrations of these drugs have been reported to inhibit lymphocyte capping in similar conditions¹⁵. Of the five drugs shown in Fig. 1, four are known to cause congenital malformations *in vivo*¹⁶. We have been unable to determine whether lidocaine is a teratogen.

Curves similar to those of Fig. 1 have been generated for 55 drugs with known teratogenic properties. To characterise the inhibitory activity of a drug we have found it useful to determine the ID₅₀—the dose of drug that inhibits 50% of tumour cell attachment. Table 1 lists these 55 drugs according to their inhibitory properties. Of drugs found to inhibit attachment, 31 of 33 (94%) have been reported to be teratogenic *in vivo*. Note the wide range of agents which inhibit attachment; these include

Table 1 Inhibition of tumour cell attachment to concanavalin A-derivatised polyethylene disks

Inhibitory drug				Non-inhibitory drug			
	ID ₅₀	Teratogen	Ref.		Maximum concentration	Teratogen	Ref.
Aspirin	14 mM	+	16	Colchicine*	5 mM	+	16
Phenylbutazone	0.12 mM	+	17	Chloroform	33.5 mM	–	18
Procaine	11 mM	+	16	Heparin	5 U ml ⁻¹	–	16
Warfarin	0.8 mM	+	16				
Caffeine	35 mM	+	16				
Diphenylhydantoin	0.84 mM	+	16				
Diazepam	0.88 mM	+	16				
Meprobamate	2.3 mM	+	16				
Chlordiazepoxide	0.17 mM	+	16				
Trifluoperazine	0.1 mM	+	16	Thalidomide*	1 mM	+	16
Chlorpromazine	0.1 mM	+	16				
Pentobarbital	1.4 mM	+	16	Hydrocortisone-21-hemisuccinate*	1 mM	–	16
Dexamethasone	0.2 mM	+	16	Cortisone acetate	0.3 mM	+	16
Prednisolone-21 sodium succinate	12 mM	+	16	Lysine	45 mM	–	16
Phenylalanine	90 mM	+	16	Cysteine	75 mM	–	16
Retinoic acid*	0.4 mM	+	16	Thiamine	2.5 mM	–	16
				Ascorbic acid	75 mM	–	16
				Ouabain	4.5 mM	–	16
Cytochalasin B*	0.69 mM	+	23	Ammonium chloride	75 mM	–	16
Dinitrophenol	0.1 mM	+	16	EDTA	10 mM	–	16
DMSO	0.9 M	+	16	Monosodium glutamate	150 mM	–	19, 20
Tween	0.38 mg ml ⁻¹	–	16	Propylene glycol	0.1%	–	16
Linoleic acid	0.053 mM	–	24	Carboxymethyl cellulose	9 mg ml ⁻¹	–	16
				Trypan blue	0.83 mM	+	21
Penicillin G	17 mM	+	22	Streptomycin	16 mM	–	25
Chloramphenicol	2.8 mM	+	16				
Nalidixic acid	0.5 mM	+	16				
Triton X-100	0.015 mM	+	16				
Ethanol	0.54 M	+	16				
Busulfan	1.3 mM	+	16	Actinomycin D	0.2 mM	+	16
BCNU†	2.3 mM	+	16	Methotrexate	2.2 mM	+	16
				5-Fluorouracil	1.9 mM	+	16
				Vinblastine	1.1 mM	+	16
Griseofulvin	0.14 mM	+	16				
Diethylstilboestrol†	0.11 mM	+	16				
Oestradiol†	0.11 mM	+	16				
Adrenaline HCl	0.46 mM	+	16				
Noradrenaline HCl	0.24 mM	+	16				

Drugs were tested for inhibition of attachment as described in the text and in Fig. 1 legend. In nearly all cases incubation with drugs was for 30 min at 37 °C. The assignment of *in vivo* teratogenic activity was based on the reference cited. +, Teratogenic *in vivo* according to cited ref.; –, non-teratogenic, according to cited ref. The maximum concentration tested was limited by solubility of the drug in phosphate-buffered saline or by its toxicity as judged by trypan blue exclusion.

* Drug dissolved in dimethyl sulphoxide (DMSO). The final concentration of DMSO did not exceed 0.3 M, a non-inhibitory concentration.

† Drug dissolved in ethanol. The final ethanol concentration was not inhibitory.

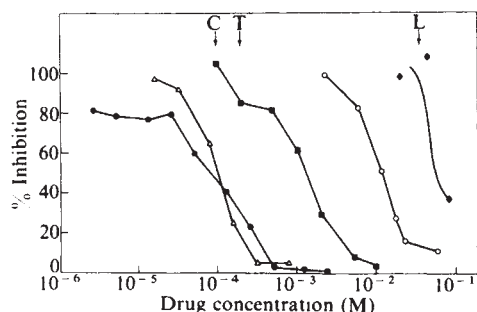


Fig. 1 Inhibition of MOT cell attachment to concanavalin A-derivatised disks. Disks were cut from 15-ml polyethylene with a 1.25 cm diameter die and derivatised in a solution of 0.1 mg ml⁻¹ concanavalin A and 2.5% glutaraldehyde overnight. The disks were washed in buffer and stored in 0.3 M glycine. Just before use they were again washed in buffer. The surface density of concanavalin A receptors on disks prepared in this manner has been estimated by tritium-labelled glucose binding to be 3×10^{12} receptors cm⁻². In this experiment ³H-labelled MOT cells at 5×10^6 per ml were incubated with the indicated concentration of drug for 15 min at room temperature and attachment to the disks determined as discussed in the text. Triplicate samples were assayed, averaged and the data expressed in terms of attachment of untreated cells. ●, Chlorpromazine (Thorazine); △, trifluoperazine (Stelazine); ■, pentobarbital (Nembutal); ○, procaine (Novocain); ◆, lidocaine (Xylocaine). The labelled arrows indicate the concentrations of chlorpromazine (C), trifluoperazine (T) and lidocaine (L) required to inhibit IgG-mediated lymphocyte capping¹⁵.

antipsychotics, anti-inflammatory agents, oestrogens and steroids, fatty acids, anaesthetics, tranquilisers, an anticonvulsant, an alkylating agent, an antifungal antibiotic, an anticoagulant, detergents and an alcohol.

Although inhibition of attachment corresponds to teratogenic activity *in vivo*, failure to inhibit does not correlate as well with lack of teratogenic activity. Table 1 indicates that of 22 drugs found to be non-inhibitory, 14 (64%) were also non-teratogenic *in vivo*. The remaining 8 (36%) thus constitute false negatives (teratogens which do not inhibit attachment). Several of these false negatives are known inhibitors of macromolecular synthesis and would not be expected to have dramatic effects on the cell surface during the 50 min of the assay. Other agents, notably Thalidomide, are believed to be teratogenic by virtue of metabolic products rather than the parent compound¹⁶. As the current assay system does not include a metabolic activation system, drugs requiring activation would not be expected to inhibit attachment. Recent experiments have shown that treatment of Thalidomide with liver microsomes yields a metabolite which inhibits attachment.

These data suggest that the inhibition of attachment *in vitro* may be useful in the prediction of teratogenic activity *in vivo*. If the five inhibitors of macromolecular synthesis and assembly are excluded, 44 of 50 tested drugs (90%) correctly predict *in vivo* teratogenic response. It is unlikely that the correspondence shown here is completely fortuitous. It is more likely that many teratogens exert their effect by interfering with normal morphological interactions and that these agents also interfere by a similar but unknown mechanism with the energy-dependent tumour cell to lectin surface interaction. However, some of these agents, such as BCNU and busulfan, are probably teratogenic because of their toxicity rather than their interference with cell to cell interactions.

Although the data reported here concern attachment of MOT cells to concanavalin A-derivatised surfaces, we have found identical attachment characteristics with cells from Ehrlich ascites tumour and with cells derived from solid tumours by collagenase digestion. Attachment to surfaces derivatised with lentil lectin or wheat-germ agglutinin was identical to that observed with concanavalin A. Attachment characteristics do not seem to depend on the nature of the plastic surface, as Teflon, Mylar, polycarbonate and acrylic surfaces derivatised with concanavalin A had characteristics identical to the polyethylene surfaces. Scanning electron micrographs of attached cells indicate that attachment is mediated exclusively by

numerous filopodia. During the 20-min attachment period, gross spreading of the tumour cells does not occur. However, there is some preliminary evidence that the filopodia extend and arborise during this interval.

If the inhibition of tumour cell attachment to lectin-coated surfaces continues to correlate well with teratogenicity, this inexpensive assay system may prove useful in the rapid identification of potentially teratogenic environmental agents and drugs.

This work was supported by Center Grant NIH CA-12662-07.

Received 7 May; accepted 28 September 1979.

1. Saxen, L. *Handbook of Teratology* Vol. 2 (eds Wilson, J. G. & Fraser, F. C.) 171-197 (Plenum, New York, 1977).
2. Wartiovaara, J., Nordling, S., Lehtonen, E. & Saxen, L. *J. Embryol. exp. Morph.* **31**, 667-682 (1974).
3. Chase, H. B. & Chase, E. B. *J. Morph.* **68**, 279-301 (1941).
4. McGuire, E. J. & Burdick, C. L. *J. Cell Biol.* **68**, 80-89 (1976).
5. Walther, B. J., Ohman, R. & Roseman, S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1569-1573 (1973).
6. Umbreit, J. & Roseman, S. *J. biol. Chem.* **250**, 9360-9368 (1975).
7. Gottlieb, D. I. & Glaser, L. *Biochem. biophys. Res. Commun.* **63**, 815-821 (1975).
8. Rutishauser, U., Thiery, J., Brackenbury, R., Sela, B. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 577-581 (1976).
9. Barbera, A. J., Marchase, R. B. & Roth, S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2482-2486 (1973).
10. Urie, J. *Cancer* Vol. 3 (ed. Becker, F. F.) 21-55 (Plenum, New York, 1975).
11. Rutishauser, U. & Sachs, L. *J. Cell Biol.* **66**, 76-85 (1975).
12. Killian, J. J. & Kollmorgen, G. M. *Nature* **259**, 674-676 (1976).
13. Kinzel, V., Kubler, D., Richards, J. & Stohr, M. *Science* **192**, 487-489 (1976).
14. Order, S. E., Donahue, V. & Knapp, R. *Cancer* **32**, 573-579 (1973).
15. Ryan, G. B., Unanue, E. R. & Karnovsky, M. J. *Nature* **250**, 56-57 (1974).
16. Schardein, J. L. *Drugs as Teratogens* (CRC Press, Cleveland, Ohio, 1976).
17. Larzen, V. & Bredahl, E. *Acta pharmac. tox.* **24**, 433-455 (1966).
18. Thomas, D. J., Warner, S. D. & Robinson, V. B. *Tox. appl. Pharmac.* **29**, 348-357 (1974).
19. Ebert, A. G. *Tox. appl. Pharmac.* **17**, 274 (1970).
20. Semprini, M. E., D'Amico, A. & Mariani, A. *Nutr. Metab.* **16**, 276-284 (1974).
21. Momie, I. W., Takacs, E. & Warkany, J. *Anat. Rec.* **156**, 175-190 (1966).
22. Filippi, B. *Adv. Teratol.* **2**, 239-256 (1967).
23. Linville, G. & Shepard, T. H. *Clin. Res.* **20**, 281 (1972).
24. Cutler, M. & Schneider, R. *Fd Cosmet. Tox.* **11**, 935-942 (1973).
25. Edelman, G. M. & Rutishauser, U. *Meth. Enzym.* **34**, 195-225 (1974).

Antibody-mediated enhancement of Flavivirus replication in macrophage-like cell lines

J. S. M. Peiris & J. S. Porterfield

Sir William Dunn School of Pathology, University of Oxford, UK

Interactions between animal viruses and antiviral antisera may exceptionally result in an apparent increase in viral infectivity¹⁻³. Halstead and coworkers⁴⁻⁶ demonstrated enhanced replication of dengue virus (a Flavivirus, family *Togaviridae*) in human or simian peripheral blood leucocytes carrying Fc receptors at subneutralising concentrations of anti-dengue antibody. We have used three continuous cell lines⁷⁻⁹ which express macrophage markers to explore the mechanism of this phenomenon. Dengue virus failed to replicate in these cells, but West Nile virus, another Flavivirus, replicated in all three, and we were able to demonstrate reproducibly 50-100-fold enhancement of virus yields in the presence of Flavivirus antisera, the effect also being directly demonstrable in P388D₁ cells by increased numbers of virus-induced plaques. The phenomenon of antibody-dependent enhancement of viral replication is not unique to dengue virus, and may have far wider relevance in other viral infections.

P388D₁ cells were used⁷. This was continuous line derived from a methyl cholanthrene-induced mouse tumour, which shows a number of macrophage characteristics, for example macrophage-like morphology, immune phagocytosis, Fc and C3 receptors, production of lysozyme, effector functions in antibody-dependent cell mediated cytotoxicity assays, and the absence of surface immunoglobulin^{10,11}. Also a second mouse cell line, J774.1 (ref. 8) and a continuous human histiocytic lymphoma cell line, U937 (ref. 9).

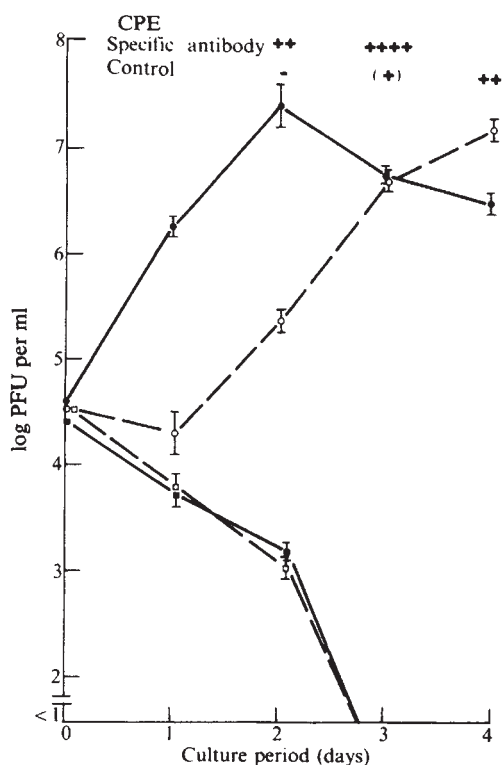


Fig. 1 Effect of specific antibody on West Nile virus yields in P388D₁ cells. P388D₁ cells (1.3×10^6) in 60 mm Falcon tissue culture dishes were infected with West Nile virus (multiplicity of infection 0.02). Rabbit antibody to normal mouse brain or West Nile virus infected mouse brain at a final dilution of 1:1,000 was added to respective culture dishes. Aliquots of supernatant medium were removed and titrated for virus infectivity in PS cells at days 0, 1, 2 and 3 post-infection. Control tissue culture dishes without cells were set up to estimate thermal inactivation of virus. Each point represents the mean of six titrations, and the 95% confidence limits of the mean is denoted. Infectious virus titre in P388D₁ cells in presence of antibody to: ○—○ normal mouse brain; ●—● WNV-infected mouse brain. Thermal inactivation of West Nile virus (identical experiment carried out in the absence of P388D₁ cells) in presence of antibody to: □—□ normal mouse brain; ■—■ WNV-infected mouse brain. The degree of cytopathic effect in the P388D₁ cell monolayer is represented: —no CPE; (+) less than 10% of cells affected; +10–25% cells affected; ++ 25–50% cells affected; +++ 50–75% cells affected; ++++ 75–100% cells affected.

P388D₁ cells were grown in MEM-Alpha medium (Gibco-Bio-cult) with 10% heat inactivated fetal calf serum (Sera-Lab) and antibiotics. West Nile virus (WNV) (Egypt 101) stock was a 10% homogenate of infected suckling mouse brain stored in aliquots at -70°C (infectious titre 2×10^7 plaque forming units (PFU) per ml). A 28-day rabbit antiserum raised against WNV was used in experiments where antibody was added. Titrations

of virus infectivity were done using pig kidney PS cells¹². Figure 1 shows the infectious virus yield from P388D₁ cells infected with WNV at a multiplicity of infection of 0.02; in the presence of anti-WNV antibody at subneutralising concentrations this was increased 100-fold in comparison with control cultures. The virus yield from antibody enhanced cultures of P388D₁ cells peaks on the second day post-infection, when there were no viable cells left to support viral replication. Yields from control cultures attain the same final virus titre as 'antibody-enhanced' culture plates, but about two days later.

In the presence of antibody, marked cytopathic effect was observed on the second post-infection day which progressed to total destruction of the cell sheet by day 3. In contrast, culture dishes infected in identical conditions without antibody, showed a mild cytopathic effect on day 3 which progressed to total cell destruction on post-infection day 5. Control experiments showed that the effect of the anti-WNV serum was not simply to protect WNV from thermal inactivation. Both the 'enhancement' of replication and the earlier cytopathic effect caused by anti-WNV antibody were reproducible even when unabsorbed virus was removed after the initial infection by two cycles of washing in phosphate-buffered saline.

We next demonstrated antibody-mediated enhancement of WNV plaque formation in P388D₁ cells under carboxymethyl cellulose overlay. Figure 2 shows the dramatic increase in the number of plaques resulting from the inclusion of specific antiserum at subneutralising concentrations (final dilution 10^{-3}) in the culture medium. This technique was used to study the ability of a range of antisera at different dilutions to demonstrate antibody-mediated enhancement of plaque formation by WNV (Table 1). Using the anti-WNV serum, it was noted that high concentrations of antibody (up to a final dilution of 10^{-2}) result in neutralisation of plaque formation. At subneutralising concentrations (final dilution 10^{-3} – 10^{-4}) a 30–40-fold increase in plaque number occurs, and at higher dilutions of antiserum, the plaque counts approach control numbers. Appropriate concentrations of antibody against other relatively closely related (dengue) or more distantly related (Kyasanur Forest disease) Flaviviruses were capable of enhancing WNV, but antisera against the antigenically unrelated Alphavirus, Semliki Forest virus, or against normal mouse brain did not affect the plaque count. This suggests that antibodies to group-specific antigenic determinants are capable of producing enhancement.

Experiments with the cell lines J774.1 and U937 infected with WNV confirm the increased yield shown in P388D₁ cells in the presence of specific antiserum, but plaque formation was not possible in these two cell lines. In contrast, virus yields from PS cells (an epithelial cell line of pig kidney origin lacking Fc receptors¹³) were not significantly altered by the presence of 'enhancing' concentrations of antibody (unpublished observations). However, plaque formation by WNV in PS cells seems to be marginally increased by appropriate concentrations of antibody (1.7-fold over control plaque members $P < 0.001$ —unpublished observation). In experiments with human peripheral blood leukocytes, WNV shows antibody mediated

Table 1 Enhancement of West Nile virus plaque formation by homologous, cross reacting and non-cross reacting antisera.

Antiserum	'Enhancement factor' at different antibody dilutions					Control	Antibody titre vs. West Nile virus	
	1/25	1/100	1/1000	1/10,000	1/100,000		H.A.I.	PRNT
Pre immune serum	1.4	0.7	1.1	0.7	1.1	1	<10	<10
Anti West Nile immune serum	0	1.0*	20	18	1.2	1	320	64
Anti Dengue type 2 immune serum	12	19	0.7	0.8	1.1	1	20	<10
Anti Kyasanur forest disease immune serum	9	2	0.9	1.2	0.8	1	40	<10
Anti Semliki Forest virus immune serum	1.1	0.8	0.9	0.9	1.2	1	<10	<10

The technique used is as for Fig. 2. The numbers represent the 'enhancement factor' at a given dilution of antibody as calculated by mean PFU in presence of antibody dilution/mean PFU in absence of antibody. The virus input was adjusted to give 5–10 PFU in control plates (viz in the absence of antibody). * Plaques were smaller than control. HAI = haemagglutination-inhibition antibody titre. PRNT = plaque reduction neutralisation titre.

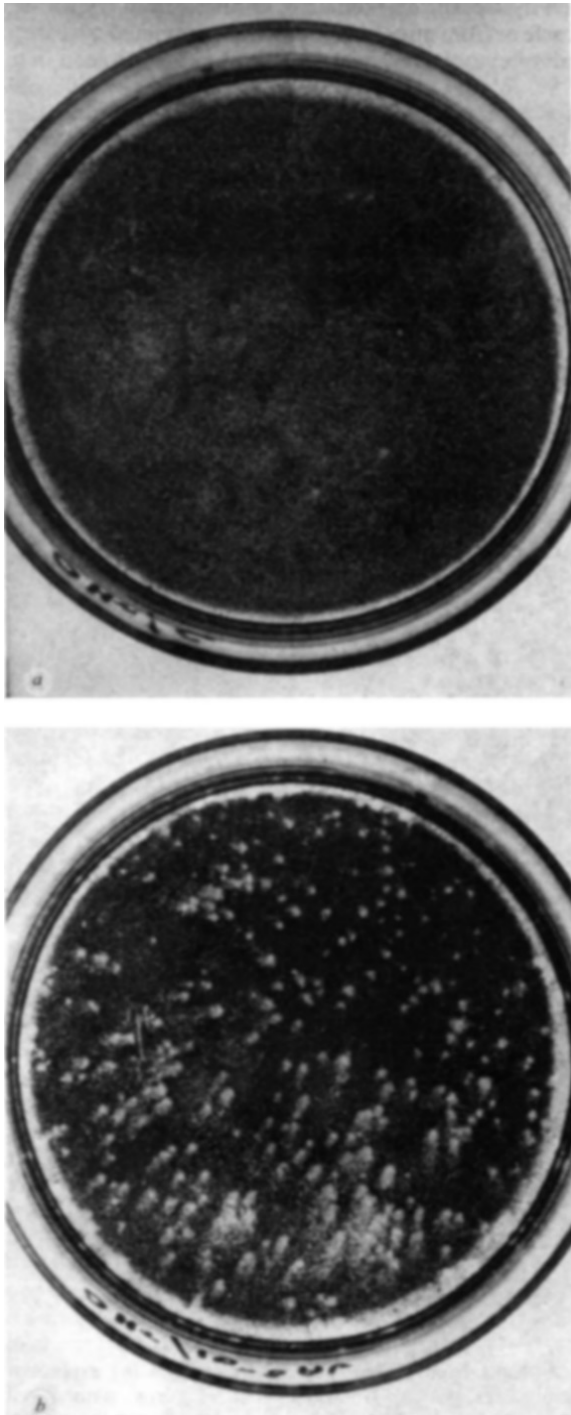


Fig. 2 Enhancement of West Nile plaque formation in P388D₁ cells by homologous antibody. 1 ml of virus dilution was mixed with an equal volume of *a*, medium; or *b*, specific antibody at 1:1,000 final dilution in 60-mm tissue culture dishes (Falcon). P388D₁ cells (1 ml) (1.8×10^6 cells ml⁻¹) was added and culture dishes incubated for 4 h in a CO₂ incubator at 35 °C. Overlay medium (3 ml) containing carboxymethylcellulose (1.5%) was added and culture dishes reincubated for 3 days. At this time the cell monolayer was stained with 1% naphthalene black.

enhancement of virus replication (unpublished observation) thus suggesting that dengue virus is not unique in this respect.

The mechanism of antibody mediated enhancement of viral replication is still open to speculation. The work of Hawkes¹, Halstead and coworkers^{4,5} and our own, has been done on cell populations containing mononuclear phagocytic cells (chick embryo fibroblasts have been shown to contain large numbers of phagocytic cells⁶). Fc receptors on such cells may promote the entry of infectious immune complexes formed in the presence of

subneutralising concentrations of antibody. Whether antibody directed against a particular antigenic determinant of Flaviviruses is capable of both 'enhancement' and neutralisation depending upon concentration, or whether antibodies of certain specificities are incapable of neutralisation and form 'infectious immune complexes' is uncertain.

The universality of this phenomenon among other virus groups remains to be investigated. Preliminary experiments indicate that antibody also enhances Alphavirus replication in the P388D₁ cells (unpublished observations). We believe that the experimental model described could contribute to our understanding of the pathogenesis of diseases such as dengue haemorrhagic fever and other viral infections.

We thank Dr Siamon Gordon for discussions and for providing the macrophage-like cell lines used in this study and Julie Wickson and Alan Hayle for technical assistance. J.S.M.P. was the recipient of a Commonwealth Tropical Medicine Research Grant.

Received 30 July; accepted 28 September 1979.

1. Hawkes, R. *Aust. J. exp. Biol. med. Sci.* **42**, 465-482 (1964).
2. Della-Porta, A. J. & Westaway, E. G. *J. gen. Virol.* **38**, 1-19 (1978).
3. Clerx, J. P. M., Horzinek, M. C. & Osterhaus, A. D. M. E. *J. gen. Virol.* **40**, 297-308 (1978).
4. Halstead, S. B. & O'Rourke, E. J. *Nature* **265**, 739-741 (1977).
5. Halstead, S. B. & O'Rourke, E. J. *J. exp. Med.* **146**, 201-217 (1977).
6. Halstead, S. B., O'Rourke, E. J. & Allison, A. C. *J. exp. Med.* **146**, 218-229 (1977).
7. Koren, H. S., Handwerker, B. S. & Wunderlich, J. R. *J. Immun.* **114**, 894-897 (1975).
8. Ralph, P., Prichard, J. & Cohn, M. *J. Immun.* **114**, 898-905 (1975).
9. Sundstrom, C. & Nilsson, K. *Int. J. Cancer* **17**, 565-577 (1976).
10. Unkeless, J. C. & Eisen, H. N. *J. exp. Med.* **142**, 1520-1533 (1975).
11. Snyderman, R., Pike, M. C., Fisher, D. G. & Koren, H. S. *J. Immun.* **119**, 2060-2066 (1977).
12. Madrid, A. T. de & Porterfield, J. S. *Bull. Wld Hlth Org.* **40**, 113-121 (1969).
13. Inoue, U. K. & Ogura, R. *Virology* **16**, 205-207 (1962).

Microtubule organisation in fibroblasts from dystrophic chickens and persons with Duchenne muscular dystrophy

Joe A. Connolly*, Vitauts I. Kalnins* & Brian H. Barber†

* Division of Histology, Department of Anatomy and

† Department of Microbiology and Parasitology, University of Toronto, Canada M5S 1A8

In spite of many attempts to correlate the expression of the Duchenne muscular dystrophy (DMD) genotype with defined phenotypic changes in both muscle and non-muscle cells from affected individuals¹⁻⁴, the molecular basis of this X-linked genetic defect has not yet been identified. However, Shay and Fuseler reported recently⁵ that interphase cells derived from the muscle of newborn dystrophic chickens had greatly diminished cytoplasmic microtubule complexes, and they speculated that this might be due to a specific defect in the microtubule-organising centres. Insofar as microtubules are a major component of the intracellular cytoskeleton⁶, a network thought to interact with the cell-surface membrane⁷, the possibility that a microtubule defect might explain the membrane alterations associated with muscular dystrophies¹⁻⁴ makes this finding extremely interesting. To assess the generality of this finding with respect to chicken dystrophy and to extend the approach to DMD, we used indirect immunofluorescence with an antiserum to purified brain tubulin to examine the organisation of microtubules in cells from the muscle of dystrophic chickens and in skin fibroblasts from persons with DMD. Unlike Shay and Fuseler⁵, we find no lack of a microtubular network in cells derived from cardiac or skeletal muscle of chickens carrying the dystrophic gene (*am*). Furthermore, our examination of fibroblasts from persons with DMD revealed a similar extensive network of microtubules in each case.

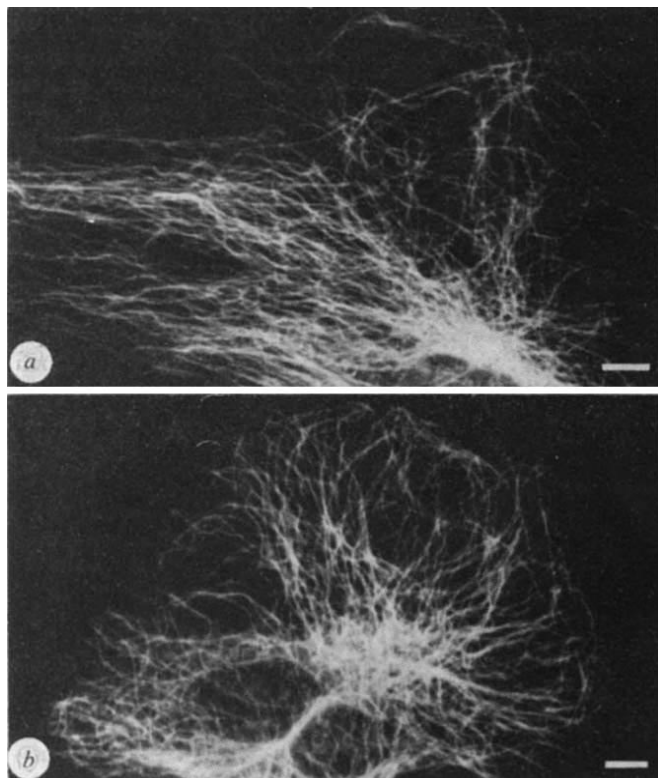


Fig. 1 Fibroblasts prepared by trypsin digestion from 1-d-old dystrophic (a) and normal (b) chicken muscle stained with antiserum to tubulin. In both cell types a similar cytoplasmic network of microtubules is clearly evident. Scale bar, 5 μ m.

When dystrophic chicken cells, cultured from trypsin-digested muscle or from muscle explants were examined, an extensive network of cytoplasmic microtubules was always seen in more than 98% of the cells (Table 1). Figure 1 shows fibroblasts prepared by trypsin digestion of skeletal muscle from 1-d-old birds. In both dystrophic (Fig. 1a) and normal (Fig. 1b) cells, the network of cytoplasmic microtubules is clearly evident. Figure 2 shows portions of two cells from explant cultures of dystrophic (Fig. 2a) and normal (Fig. 2b) skeletal muscle cultured from 15-d-old birds. Again, no obvious differences in the organisation of microtubules were evident. Similar results were obtained independent of the age of the bird from which the cells were cultured, whether skeletal or cardiac muscle was used as the source of cells in explant cultures, and with both fixation procedures. Similar results were also obtained when cells cultured from control birds were examined. These findings are summarised in Table 1. Only very faint diffuse staining was seen when preimmune serum from the same rabbit was used, or when the fluorescein-labelled goat IgG alone was used.

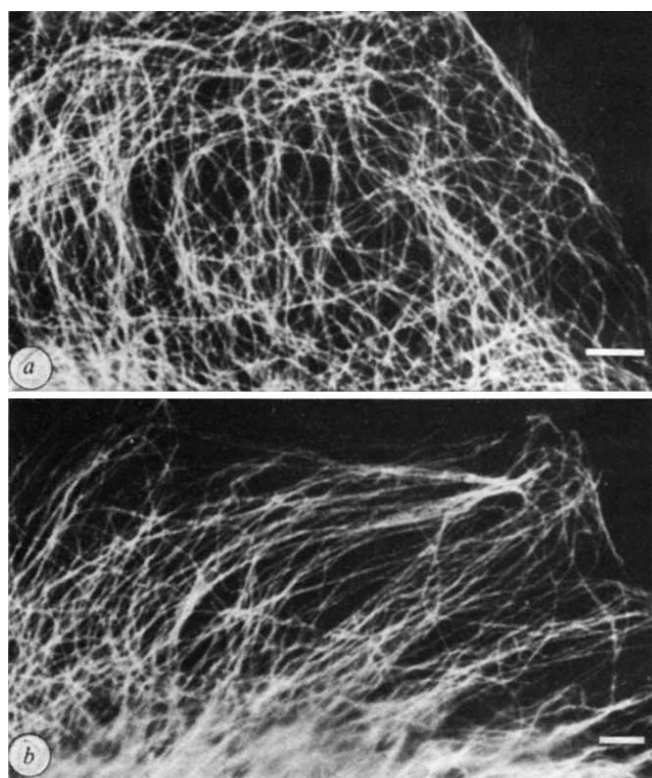


Fig. 2 Cells from dystrophic (a) and normal (b) explants of skeletal (leg) muscle from 15-d-old chickens stained with antiserum to tubulin. No differences are apparent in these higher magnification views of the microtubular network. Scale bar, 5 μ m.

To obtain dystrophic chicken cells, we used chickens of the Connecticut line⁸. These birds carry the dystrophic gene (*am*) and are characterised phenotypically by the onset of clinical symptoms 2–3 weeks after hatching⁹. White Leghorn chickens were used as controls. Cells were prepared in one of two ways. (1) A single cell suspension was prepared from trypsinised skeletal muscle tissue⁹ and grown to confluency in α -minimal essential medium (MEM) containing 10% fetal bovine serum (FBS). Cells were examined after a single subculture in the same medium. (2) Small pieces of cardiac and skeletal (leg) muscle were placed on 22-mm glass coverslips, allowed to attach for 10 min, and then 2 ml of either α -MEM or CMRL 1066 (GIBCO) culture medium containing 10% horse serum and 5% FBS was added, according to the method of Shay and Fuseler⁵. Cells from these muscle explant cultures were fixed for immunofluorescence after 7 d in culture. Fibroblasts were established in culture from skin biopsies of persons with DMD (three males, 4–5 yr old) and controls of the same sex and approximate age (two males, 4 and 6 yr old). Cells were recovered from storage in liquid nitrogen, and grown at 37 °C under 5% CO₂ in α -MEM supplemented with 15% FBS and antibiotics (100 U per ml penicillin, 100 μ g per ml streptomycin). Cells were subcultured routinely every 2–3 weeks and examined by immunofluorescence between the fifth and ninth subculture.

For immunofluorescence, cells were fixed either with methanol followed by acetone at –20 °C (refs 10 and 11), or with methanol alone after treatment with 0.5% Triton X-100 (ref. 12). Fixed cells were incubated with a previously characterised rabbit antiserum to electrophoretically purified pig brain tubulin^{10,11} diluted 1:30 in phosphate-buffered saline (PBS) for 40 min at room temperature, washed three times in PBS, and incubated for 30 min with a 1:5 dilution of fluorescein-conjugated goat IgG directed against rabbit IgG (Hyland Diagnostics). Cells were then washed three times in PBS and mounted in 50% glycerol in PBS, pH 7.8.

In a complementary study, we also examined microtubule organisation in skin fibroblasts of persons with DMD. In cultured fibroblasts from three individuals with DMD and two controls, no differences were seen between normal and dystrophic cells, and in all cases an extensive network of microtubules was visible in more than 99% of the cells. Figure 3 shows the clearly evident microtubular network present in cultured dystrophic (Fig. 3a) and normal (Fig. 3b) human skin fibroblasts. Again, similar results were obtained using either fixation method.

We were unable to detect significant differences in the organisation of the cytoplasmic network of microtubules between cells derived from dystrophic chicken muscle and control cells. Likewise, we were unable to detect differences in microtubule organisation between cultured fibroblasts from

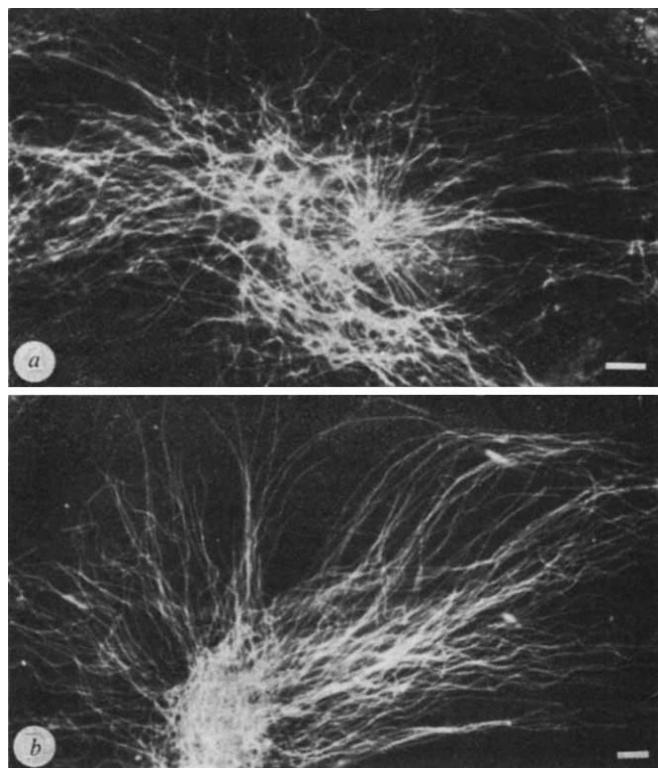


Fig. 3 Human skin fibroblasts from a person with DMD (a) and a normal control (b) stained with antiserum to tubulin. A similar cytoplasmic network of microtubules is present in both cells. Scale bar, 5 μ m.

persons with DMD and control subjects. These results contrast with those reported by Shay and Fuseler⁵. The methods we used for explant cultures were identical to those described in their report⁵. Like Shay and Fuseler, we observed that both normal and dystrophic interphase cells are flat and well-spread on the substrate, and that the distribution of actin in these two cell types is similar (B.H.B., unpublished results). Therefore, we do not understand the discrepancy between their results and ours with

respect to the distribution of microtubules in dystrophic cells. The dystrophic chickens from which we prepared our cells were of the Connecticut line, whereas Shay and Fuseler used cells from line 413 birds⁵. Although the difference in genetic background could contribute to the observed differences, it seems unlikely to represent a major factor because both chicken lines are derived from the original 301 dystrophic line, possess the same dystrophic gene (*am*) and are characterised by similar manifestations of the disease⁸. Differences between the antiserum used by Shay and Fuseler and our antiserum may also account in part for the discrepancy in results. However, we believe a more likely explanation lies in the different fixation procedures used and that our methods enhanced the visualisation of microtubules by reducing background staining.

In conclusion, we have been unable to detect major differences in the organisation of microtubules between cells from dystrophic individuals and their corresponding control cells. On the basis of these findings, therefore, we are unable to support any correlation between the expression of the dystrophic genotype and a reduction in the number of microtubules in the interphase cells of affected individuals.

This work was supported by grants from the Ontario Heart Foundation and the MRC of Canada to V.I.K. and from the Muscular Dystrophy Association of Canada to B.H.B. We thank Dr Pat Stewart for dystrophic and normal chickens, Dr M. Thompson for establishing and making available human fibroblast strains, and Dr L. Pierro for maintenance of the chicken cell line. J.A.C. is the recipient of an MRC studentship.

Received 30 July; accepted 28 September 1979.

- Engle, W. K. *Adv. Neurol.* **17**, 197–226 (1977).
- Schotland, D. L., Bonilla, E. & Van Meter, M. *Science* **196**, 1005–1007 (1977).
- Kobayashi, T., Mawatari, S. & Kuroiwa, Y. *Clin. chim. Acta* **85**, 259–266 (1978).
- Pickard, N. A. *et al. New Engl. J. Med.* **299**, 841–846 (1978).
- Shay, J. W. & Fuseler, J. W. *Nature* **278**, 178–180 (1979).
- Osborn, M., Born, T., Koitsch, H.-J. & Weber, K. *Cell* **14**, 477–488 (1978).
- Nicolson, G. L. *Biochim. biophys. Acta* **457**, 57–108 (1976).
- Wilson, B. W., Randall, W. R., Patterson, G. T. & Entrikin, R. K. *Ann. N.Y. Acad. Sci.* **317**, 224–246 (1979).
- Moscona, A. *Expl Cell Res.* **3**, 535–539 (1952).
- Connolly, J. A. & Kalnins, V. I. *J. Cell Biol.* **79**, 526–532 (1978).
- Connolly, J. A., Kalnins, V. I., Cleveland, D. W. & Kirschner, M. W. *J. Cell Biol.* **76**, 781–786 (1978).
- Osborn, M. & Weber, K. *Cell* **12**, 561–571 (1977).

Table 1 Cytoplasmic microtubules in chicken cells

Cell type	Age of animal	% Positive* cells
E-NS	12d (e)	99.5
E-NC	12d (e)	99.0
E-DS	12d (e)	99.5
E-DC	12d (e)	99.5
E-NS	1d (h)	98.5
E-NC	1d (h)	99.0
E-DS	1d (h)	98.5
E-DC	1d (h)	98.5
E-NS	15d (h)	100.0
E-NC	15d (h)	99.0
E-DS	15d (h)	98.5
E-DC	15d (h)	98.5
T-NS	12d (e)	100.0
T-DS	12d (e)	100.0
T-NS	1d (h)	100.0
T-DS	1d (h)	100.0

E, cells from muscle explant cultures; T, cells from trypsinised muscle tissue; NS, normal skeletal muscle; NC, normal cardiac muscle; DS, dystrophic skeletal muscle; DC, dystrophic cardiac muscle; e, embryonic; h, hatchling.

* In each case 200 cells were counted and cells were considered positive only if numerous microtubules of extended length were visible in the cytoplasm.

Different receptors for distribution of peanut and ricin agglutinins between inner and outer segments of rod cells

C. D. B. Bridges & Shao-Ling Fong

Cullen Eye Institute, Baylor College of Medicine, Houston, Texas 77030

Lectins can be used as probes for cell-surface oligosaccharides^{1–4}. These proteins display high specificities for certain haptene sugars, although the details of the sugar linkages and the three-dimensional array of the oligosaccharide may all be involved in determining the affinity of a lectin for its receptor. We have now shown that peanut and ricin agglutinins bind differentially to the surfaces of rod inner and outer segments.

Receptors for jack bean lectin (concanavalin A, Con A) are present on the surfaces of rod outer segments (ROS)^{5–7}. Con A has a specificity for complex mannose-containing oligosaccharides with internal and terminal *N*-acetylglucosamine residues^{8–11} as well as simpler structures containing mannosyl units¹². The high affinity of Con A for rhodopsin¹³ is consistent with recent observations showing that its oligosaccharide has the structure GlcNAc β 1 $\rightarrow$ 2Man α 1 $\rightarrow$ 3(Man α 1 $\rightarrow$ 6)Man β 1 $\rightarrow$ 4GlcNAc β 1 $\rightarrow$ 4GlcNAc $\rightarrow$ Asn (refs 14, 15). Because rhodopsin is probably present in the plasma membrane of the ROS as well

as the bimembranous disks found in its interior^{16,17}, it may represent the majority of the Con A receptors. A minor ROS protein, with a molecular weight of 291,000 (refs 18, 19), has also been shown to bind Con A²⁰, and probably does not play a significant part in binding Con A to the ROS surface.

Some rhodopsin molecules may carry terminal D-galactose residues²¹, so we also investigated the galactose-specific agglutinins from *Ricinus communis* (RCA_I) and *Arachis hypogaea* (PNA). These lectins have been used to suggest the existence of terminal galactose units in, for example, normal 3T3, transformed SV3T3, transformed 3T12 cells and embryonal carcinoma cells^{22,23}.

PNA (L'Industrie Biologique Française) and RCA_I (Sigma and PL-Biochemicals) caused agglutination in suspensions of washed ROS at concentrations of 11 and 25 $\mu\text{g ml}^{-1}$, respectively. The effect was inhibited by D-galactose at a concentration of 4–40 mM for PNA and 0.3–3 mM for RCA_I. Lactose and α -D-(but not α -L-) fucose inhibited agglutination by RCA_I. Other sugars tested, and found to be ineffective, were α -methyl mannopyranoside, D-glucose, N-acetyl glucosamine, chitobiose and N-acetyl galactosamine. For comparison, Con A agglutinated ROS at 0.1–0.5 $\mu\text{g ml}^{-1}$, and was inhibited by α -methyl mannopyranoside and glucose (see ref. 5).

Fresh ROS from light- or dark-adapted frogs were incubated with fluorescein isothiocyanate (FITC)-labelled PNA or RCA_I and compared, using a Zeiss epifluorescence microscope, with ROS incubated with FITC-Con A. In such preparations a variable proportion of the ROS had part of the inner segment attached. Figure 1 is a typical fluorescence photograph of one of these organelles after incubation with FITC-Con A. Bright surface fluorescence is visible on both the outer and inner segment, showing that receptors for Con A are present at both locations. A brightly fluorescent band delineates the junction between inner and outer segment, perhaps attributable to penetration of the lectin into the few open disks known to occur at that point²⁴. The surface fluorescence was virtually abolished if incubation was carried out in the presence of 100 mM α -methyl mannopyranoside.

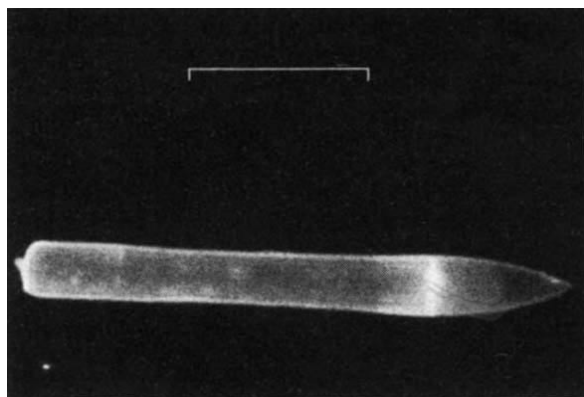


Fig. 1 Fluorescence photograph of an isolated frog (*Rana pipiens*) rod outer segment with part of inner segment attached. FITC-Con A, 100 $\mu\text{g ml}^{-1}$ in frog Ringer, 12 min incubation at room temperature. Scale bar, 20 μm .

In contrast, Fig. 2 is a similar preparation incubated with FITC-PNA. Although the fluorescence intensity is lower than that for Con A, and hence more difficult to photograph, it can be seen that the bright line of the outer segment surface label stops abruptly at the inner segment, and that there is a dividing line of similar intensity between the two sectors. The diffuse, low-level fluorescence occupying part of the inner segment is autofluorescence which develops in the region of tightly packed mitochondria. It is straw yellow in colour, in contrast to the characteristic green emission of the FITC label. RCA_I gave an

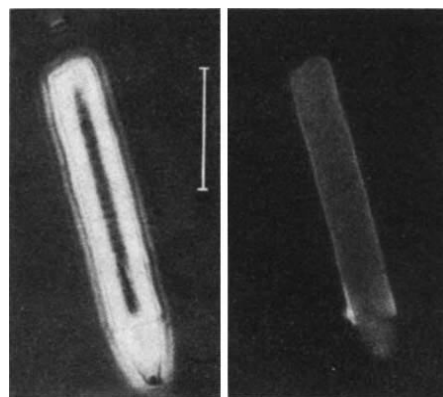


Fig. 2 Preferential binding of FITC-PNA to outer segment plasma membrane. Left hand side, normal phase illumination; right hand side, fluorescence. Both photographs were taken on Kodak Tri-X film pushed to an effective ASA rating of 1,600 with Diafine. Virtually no fluorescence is visible on the plasma membrane of the inner segment; diffuse inner segment fluorescence is autofluorescence from the mitochondria. FITC-PNA, 200 $\mu\text{g ml}^{-1}$, 12 min at room temperature. Scale bar, 20 μm .

identical pattern of labelling. The binding of both lectins was abolished in the presence of 100 mM D-galactose. After prolonged incubation with high lectin concentrations, a very faint surface label was occasionally seen on the inner segment, but in all cases this was many times weaker than that of the outer segment.

Although PNA has its highest affinity for the disaccharide D-gal β 1 $\rightarrow$ 3GalNAc (ref. 25), characteristic of many blood glycoproteins, it shares with RCA_I a specificity for the simple terminal sugar D-galactose. The difference between inner and outer segment membranes in their ability to bind RCA_I and PNA suggests that oligosaccharides present on the inner segment plasma membrane lack D-galactose in an accessible, terminal position. Rapid translational and rotational freedom for rhodopsin has been demonstrated in rod disk membranes, indicating that they have an unusually low viscosity²⁶. A similar fluidity in the plasma membrane enveloping the whole rod cell would permit RCA_I and PNA receptors to diffuse through the inner-outer segment junction. Their failure to do this may indicate the existence of a barrier at this point, or the plasma membrane may be more rigid than the disk membranes.

The synthetic machinery of the photoreceptor is found in the inner segment, but the finding that oligosaccharides binding RCA_I and PNA are restricted to in the outer segment suggests that only in this region of the cell have they been incorporated into the membrane, or their terminal sugars modified. The receptors themselves may be glycolipids or galactosylated rhodopsin molecules. As noted above, rhodopsin molecules constitute many of the Con A receptors in the outer segment plasma membrane. However, the binding observed on the inner segment (down to the synaptic pedicle in papain-dissociated material) may not involve rhodopsin because Con A binds to sugar sequences common to many glycoproteins^{27–29}.

We thank the Retina Research Foundation of Houston and the USPHS (National Eye Institute) for financial assistance.

Received 11 July; accepted 28 September 1979.

1. Lis, H. & Sharon, N. in *The Antigens* Vol. IV (ed. Sela, M.) 429–529 (1977).
2. Kornfeld, S. & Kornfeld, R. in *The Glycoconjugates* Vol. 1 (eds Horowitz, M. I. & Pigman, W.) 437–449 (Academic, New York, 1978).
3. Nicolson, G. L. *Int. Rev. Cytol.* **39**, 89–190 (1974).
4. Dulane, J. T. *Molec. cell. Biochem.* **21**, 43–63 (1978).
5. Bridges, C. D. B. *Invest. Ophthalmol. ARVO Suppl.* **156** (1978).
6. Molday, R. S. *J. supramolec. Struct.* **4**, 549–557 (1976).
7. Hall, M. O. & Nir, I. *Expl Eye Res.* **17**, 595–605 (1976).
8. Young, N. M. & Leon, M. A. *Biochim. biophys. Acta* **365**, 418–424 (1974).
9. Krusius, T., Finne, J. & Rauvala, H. *FEBS Lett.* **71**, 117–120 (1976).
10. Kornfeld, R. & Ferris, C. J. *J. biol. Chem.* **250**, 2614–2619 (1975).
11. Baenziger, J. U. & Fiete, D. *J. biol. Chem.* **254**, 2400–2407 (1979).
12. Goldstein, I. J. in *Concanavalin A as a Tool* (eds Bittiger, H. & Schnebli, H. P.) 3–15 (Wiley, London, 1976).

13. Steinemann, A. & Stryer, L. *Biochemistry* **12**, 1499–1502 (1973).
14. Fukuda, M. N., Papermaster, D. S. & Hargrave, P. A. *J. biol. Chem.* **254**, 8201–8207 (1979).
15. Liang, C.-J., Yamashita, K., Shichi, H., Muellenberg, C. G. & Kobata, A. *J. biol. Chem.* **254**, 6414–6418 (1979).
16. Jan, L. Y. & Revel, J. P. *J. Cell Biol.* **62**, 257–273 (1974).
17. Basinger, S. F., Bok, D. & Hall, M. O. *J. Cell Biol.* **69**, 29–42 (1976).
18. Dreyer, W. J., Papermaster, D. S. & Kuhn, H. *Ann. N.Y. Acad. Sci.* **195**, 61–74 (1972).
19. Bownds, D. et al. *Expl Eye Res.* **18**, 253–269 (1974).
20. Fong, S.-L. & Bridges, C. D. B. *Invest. Ophthalmol. ARVO Suppl.* **155** (1979).
21. O'Brien, P. J. in *Receptors and Recognition Series A*, Vol. 6 (eds Cuatrecasas, P. & Greaves, M. F.) 109–150 (Chapman & Hall, London, 1978).
22. Nicolson, G. L. & Lacorbiere, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1672–1676 (1973).
23. Reisner, Y. et al. *Dev Biol.* **61**, 20–27 (1977).
24. Nilsson, S. E. *J. Ultrastruct. Res.* **11**, 581–620 (1964).
25. Lotan, R., Skutelsky, E., Danon, D. & Sharon, N. *J. biol. Chem.* **250**, 8518–8523 (1975).
26. Poo, M. M. & Cone, R. A. *Nature* **247**, 438–441 (1974).
27. Zinn, A. B., Plantner, J. J. & Carlson, D. M. in *The Glycoconjugates* Vol. 1 (eds Horowitz, M. I. & Pigman, W.) 459–540 (Academic, New York, 1977).
28. Montreuil, J. *Pure appl. Chem.* **42**, 431–477 (1975).
29. Sturgess, J., Moscarello, M. & Schachter, H. in *Current Topics in Membranes and Transport* Vol. 11 (eds Bronner, F. & Kleinzeller, A.) 1–13 (Academic, New York, 1978).

Enkephalins presynaptically inhibit cholinergic transmission in sympathetic ganglia

Shiro Konishi, Akinobu Tsunoo & Masanori Otsuka

Department of Pharmacology, Faculty of Medicine, Tokyo Medical and Dental University, Yushima, Bunkyo-ku, Tokyo, Japan

Recent biochemical and immunohistochemical studies have shown that the opioid peptides, enkephalins, occur in nerve terminals and cell bodies in mammalian sympathetic ganglia^{1–3}. Opiates and enkephalins are thought to inhibit synaptic transmission in the peripheral nervous tissues as well as in the central nervous system^{4–12}. The mechanisms of the opiate actions, however, are not entirely clear; both pre- and postsynaptic sites of action have been proposed^{7–9,11,12}. As acetylcholine is known to be the major neurotransmitter in the autonomic ganglia and as the mechanism of synaptic transmission is well clarified¹³, analysis of the peptide action could be more easily but equally usefully carried out in the peripheral synapses than in central synapses. We now report that enkephalins presynaptically inhibit cholinergic transmission in sympathetic ganglia.

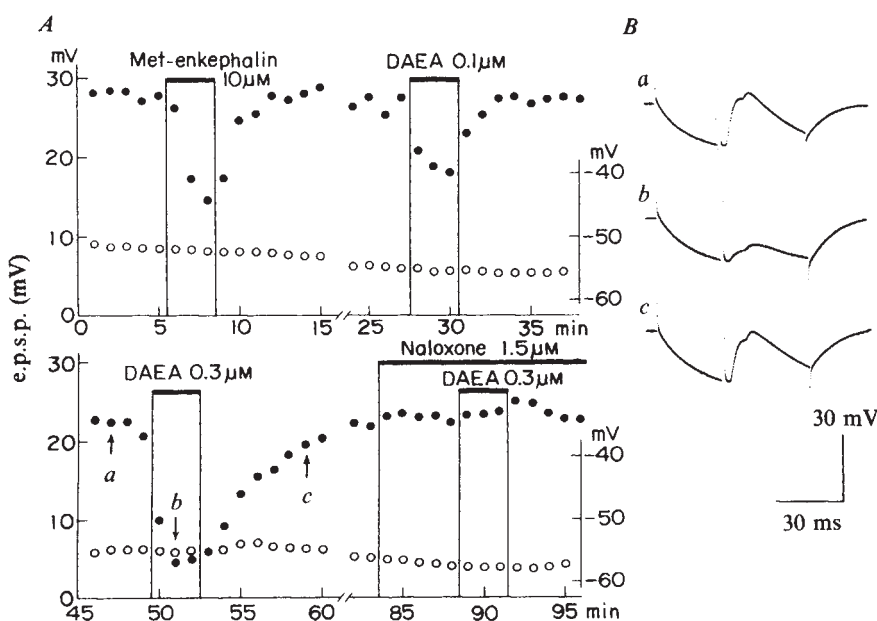


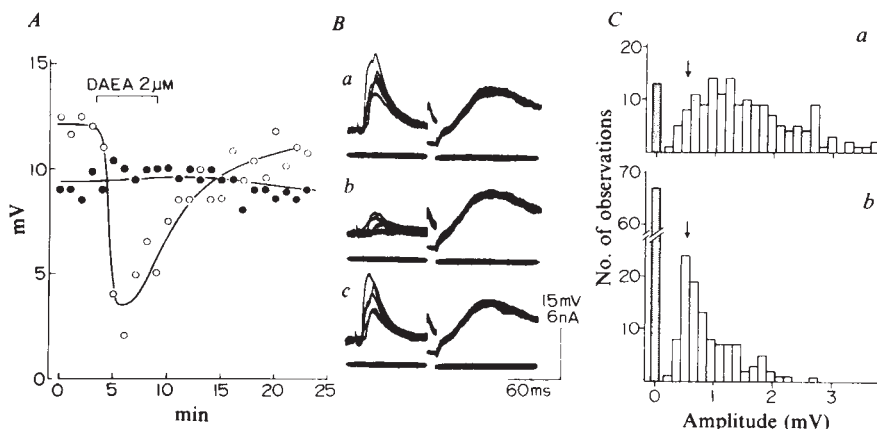
Fig. 1 Effects of Met-enkephalin and DAEA on e.p.s.ps, resting potential and membrane resistance of a postganglionic cell in the inferior mesenteric ganglion, and antagonism by naloxone. The preparation was perfused with solution containing low calcium (0.7 mM) and high magnesium (2 mM) concentrations. Hyperpolarising current pulses of 70 ms duration and 0.5 nA intensity were passed into the cell to prevent spike generation following the e.p.s.p. and also to estimate the membrane resistance. The lumbar splanchnic nerve was stimulated at 1-min intervals for 16 s at 2.5 Hz. **A**, The amplitude of the e.p.s.ps (●) was determined at 1-min intervals after averaging 32 responses. The resting potential (○) was measured from the continuous record of the pen recorder. Ordinate: left, amplitude of the averaged e.p.s.ps; right, resting potential. Each drug was perfused into the bath during the period indicated in the graph. **B**, a–c, Sample records of the averaged e.p.s.ps and the potential changes caused by current injection. Records were obtained at arrows a–c of **A**.

the amplitude of e.p.s.ps by $52.3 \pm 4.4\%$ (mean $\pm$ s.e.m., $n = 14$), without significantly affecting either the resting potential (Fig. 1A) or the membrane resistance, which was estimated by passing constant hyperpolarising current pulses (Fig. 1B) or by observing the current-voltage relationship. The inhibitory effect of DAEA (0.3 – $1.0 \mu\text{M}$) was abolished by the opiate antagonist naloxone (1 – $3 \mu\text{M}$), suggesting that the peptide inhibition results from a specific action on opiate receptors (Fig. 1A). Naloxone alone did not produce any consistent change in the size of e.p.s.ps.

number of quanta released by each stimulus. DAEA (1 – $2 \mu\text{M}$) increased the number of failures (see Fig. 2B, C) and reduced the mean quantum content of the e.p.s.ps to 40–60% of the control ($n = 3$). However, the peptide did not cause any significant change in the mean amplitude of the quantal unit (Fig. 2C). The results support the presynaptic nature of enkephalin-induced inhibition.

We also observed that noradrenaline depressed excitatory synaptic transmission in the coeliac-superior mesenteric ganglion, confirming the previous results of Christ and Nishi¹⁷.

Fig. 2 Effect of DAEA on the e.p.s.p. and acetylcholine-induced depolarisation of the postganglionic cell in the coeliac-superior mesenteric ganglion, and on amplitude histograms of e.p.s.ps. In A and B the greater splanchnic nerve was stimulated at 1-min intervals by 16 pulses at 0.5 Hz. Acetylcholine was iontophoretically applied to the cell about 80 ms after each nerve stimulus. A, The amplitudes of the e.p.s.ps (○) and the acetylcholine-induced depolarisation (●) were determined at 1-min intervals after averaging 16 responses. The peptide was perfused into the bath during the period indicated. B, a–c, Upper trace shows the e.p.s.ps (left) and acetylcholine-induced depolarisation (right), and lower trace shows the current passed into the acetylcholine-containing electrode. a, Control; b, 2–3 min after addition of DAEA ($2 \mu\text{M}$); c, 7–8 min after washing out DAEA. Hyperpolarising current (0.1 nA) was continuously passed into the cell to prevent spike generation. A and B were obtained from the same experiment. C, Amplitude histograms of e.p.s.ps evoked by stimulation of the greater splanchnic nerve at 2 Hz. a, Control; b, in DAEA ($1 \mu\text{M}$). The quantum content, m , was estimated from the numbers of stimuli, n , and failures (stippled bar). The estimated unit sizes of the e.p.s.ps are shown by arrows. For a and b, respectively, n was 167 and 176, and m was 2.55 and 0.97.



To test the possibility that enkephalins act directly on the postsynaptic membrane and reduce the sensitivity to acetylcholine, the effect of DAEA on acetylcholine response was examined. An intracellular recording electrode was inserted into superficial cells and acetylcholine was applied iontophoretically from another microelectrode placed on the surface of the impaled cells. Short pulse application of acetylcholine produced a depolarisation of the postganglionic cells. The size and time to peak of the depolarisation varied greatly depending on the position of the drug-containing electrode and the applied current. As shown in Fig. 2, the amplitude and time course of the acetylcholine-induced depolarisation were not affected by DAEA, whereas the e.p.s.ps recorded simultaneously were markedly reduced in size. Both the e.p.s.ps and the acetylcholine-induced depolarisation were almost completely abolished by dihydro- β -erythroidine ($5 \mu\text{M}$). These observations suggest that enkephalins inhibit the release of acetylcholine from the preganglionic terminals but do not reduce the sensitivity of nicotinic receptors on the postganglionic cells.

The decapeptide, substance P, which has been demonstrated in the sympathetic ganglia¹⁶, caused a depolarisation of up to 20 mV in certain postganglionic cells. The substance P-induced depolarisation was not depressed by tetrodotoxin ($1 \mu\text{M}$), which completely blocked synaptic transmission in these ganglia. This suggests that substance P acts directly on the postsynaptic membrane. As observed with the acetylcholine response, the direct depolarising action of substance P (0.4 – $1.0 \mu\text{M}$) was not affected by DAEA (1 – $2 \mu\text{M}$), suggesting that enkephalins have no effect on the chemosensitivity of the postsynaptic membrane.

To confirm the site of enkephalin action, amplitude histograms of e.p.s.ps evoked by preganglionic stimulation at 1–2 Hz in the low calcium (0.7 mM) and high magnesium (2 mM) solution were examined before and after addition of DAEA. Assuming a Poisson distribution of the e.p.s.ps and counting the number of failures in a series of trials, we estimated the mean

However, the possibility that the inhibitory effect of enkephalins might be mediated by the release of a catecholamine was excluded by the observation that the inhibition caused by DAEA ($1 \mu\text{M}$) was not affected by phentolamine ($3 \mu\text{M}$), which effectively antagonised the inhibitory effect of noradrenaline ($5 \mu\text{M}$).

Our results suggest the possible involvement of enkephalins in presynaptic inhibition of cholinergic transmission in mesenteric ganglia. This is particularly relevant to morphological evidence that some of the preganglionic fibres form axo-axonic synapses in the cat inferior mesenteric ganglion¹⁸. Furthermore, the mode of enkephalin action shown in this study is consistent with recent evidence that enkephalins depress primary afferent transmission in the spinal cord and brain stem, and inhibit the release of substance P, a putative neurotransmitter of the primary afferent fibres^{7,8,19}.

We thank Dr. T. M. Jessell for reading the manuscript and giving helpful advice.

Received 3 August; accepted 21 September 1979.

- Di Giulio, A. M. *et al. Neuropharmacology* **17**, 989–992 (1978).
- Hughes, J., Kosterlitz, H. W. & Smith, T. W. *Br. J. Pharmacol.* **61**, 639–647 (1977).
- Schultzberg, M. *et al. Neuroscience* **4**, 249–270 (1979).
- Duggan, A. W., Hall, J. G. & Headley, P. M. *Br. J. Pharmacol.* **61**, 399–408 (1977).
- Forbes, J. E. & Dewey, W. L. *Life Sci.* **19**, 401–408 (1976).
- Hughes, J., Kosterlitz, H. W. & Leslie, F. M. *Br. J. Pharmacol.* **53**, 371–381 (1975).
- Jessell, T. M. & Iversen, L. L. *Nature* **268**, 549–551 (1977).
- Mudge, A. W., Leeman, S. E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 526–530 (1979).
- North, R. A., Katayama, Y. & Williams, J. T. *Brain Res.* **165**, 67–77 (1979).
- Paton, W. D. M. *Br. J. Pharmacol.* **11**, 119–127 (1957).
- Wouters, W. & Bercken, J. V. D. *Nature* **277**, 53–54 (1979).
- Zieglansberger, W. & Tulloch, I. F. *Brain Res.* **167**, 53–64 (1979).
- Dennis, M. J., Harris, A. J. & Kuffler, S. W. *Proc. R. Soc. B* **177**, 509–539 (1971).
- Skok, U. I. *Physiology of Autonomic Ganglia*, 32–45 (Igaku Shoin, Tokyo, 1973).
- Pert, C. B., Pert, A., Chang, J.-K. & Fong, B. T. W. *Science* **194**, 330–332 (1976).
- Hökfelt, T., Elfvin, L.-G., Schultzberg, M., Goldstein, M. & Nilsson, G. *Brain Res.* **132**, 29–41 (1977).
- Christ, D. D. & Nishi, S. *J. Physiol., Lond.* **213**, 107–117 (1971).
- Elfvin, L.-G. *J. ultrastruct. Res.* **37**, 426–431 (1971).
- Otsuka, M. & Konishi, S. *Cold Spring Harb. Symp. quant. Biol.* **40**, 135–143 (1976).

ADP is a potent inhibitor of human platelet plasma membrane adenylate cyclase

Dermot M. F. Cooper & Martin Rodbell

Section on Membrane Regulation, Laboratory of Nutrition and Endocrinology, NIAMDD, NIH, Bethesda, Maryland 20205

The potent inducer of platelet aggregation, ADP, inhibits both basal and stimulated cyclic AMP production in intact platelets^{1,2}. However, several attempts have failed to detect an effect of the nucleotide on adenylate cyclase activity in broken cell preparations^{2,3}. This paradox has led to suggestions that ADP may act on unidentified sites before the adenylate cyclase step, such as a nucleoside diphosphate kinase⁴ and a Ca^{2+} -transport system^{5,6}. It is difficult to demonstrate the effect of physiologically effective concentrations (0.1–5 μM) of ADP on a crude membrane preparation of adenylate cyclase because of the presence of nucleoside phosphohydrolase activities which may degrade either the substrate ATP or the putative effector ADP to products (such as ADP, AMP and adenosine) displaying diverse effects. A substrate regenerating system which converts ADP to ATP obviously provides no solution. The contamination by adenine nucleotides and catecholamines in widely used crude lysate preparations aggravates the problem^{7–9}. Using a purified human platelet plasma membrane preparation, and the non-hydrolysable ATP analogue APP(NH)P as substrate, we have now demonstrated direct inhibition of adenylate cyclase by ADP in the physiologically effective concentration range. The inhibitory effect of ADP is immediate in onset and is equal to that produced by adrenaline; it is shared by ADP- βS (and 2'-deoxy ADP to some extent) but not by AP(CH₂)P or other adenine nucleotides. The K_i for ADP is $\sim 0.3 \mu\text{M}$ and maximal inhibition is observed at 2 μM .

In these studies membranes were prepared essentially as described by Barber and Jamieson¹⁰. The procedure involves lysis of glycerol-loaded freshly drawn washed human platelets, followed by sucrose gradient centrifugation of the particulate fraction and collection of the plasma membrane-enriched region at 60,000 g. Membranes were washed before use with 0.4 M NaCl in 20 mM Tris, pH 7.4, centrifuged at 10,000g, 10 min, 4 °C and resuspended in 10 mM Tris pH 7.4 containing 0.1% bovine serum albumin and 0.5 mM EGTA. Assays were carried out at 24 °C, generally for 6 min in an incubation volume

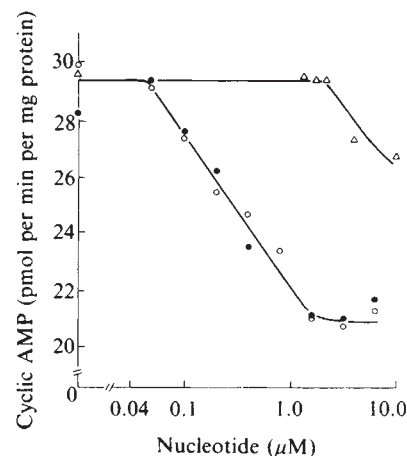


Fig. 1. Effect of ADP, ADP- βS and AP(CH₂)P on human platelet plasma membrane adenylate cyclase activity. Enzyme activity was measured in the presence of 1 μM PGE₁ and 1 μM GTP using 40 μM APP(NH)P, 2 mM MgCl₂ and 0.2 mM EGTA in the presence of the indicated concentrations of ADP (●), ADP- βS (○) and AP(CH₂)P (Δ).

of 100 μl containing 2–4 μg membrane protein, 0.04% bovine serum albumin, 0.2 mM EGTA, 30 mM Tris pH 7.4, 2 mM MgCl₂, 40 μM APP(NH)P (Böhringer, lot 31), 1 μCi [α -³²P]APP(NH)P, 50 μM cyclic AMP with other additions as indicated. ³²P-labelled cyclic AMP was isolated as described elsewhere¹¹. Triplicate or quadruplicate determinations of activity were normally performed in a block design and standard deviations of means were generally <5%. When assessed by TLC on PEI-cellulose¹², substrate concentration remained constant (>95%) during the routine assay period.

The effects of ADP, ADP- βS (Böhringer) and AP(CH₂)P (ICN) on platelet adenylate cyclase activity are compared in Fig. 1. A potent inhibition of activity (K_i : 0.3 μM) which is maximal at 2 μM is displayed by both ADP and ADP- βS . AP(CH₂)P is only weakly effective. The very similar behaviour of ADP and ADP- βS demonstrates that no phosphorylation of the nucleotide is required for its effect since the thiol analogue cannot be phosphorylated. The ineffectiveness of AP(CH₂)P correlates well with intact platelet studies which showed that this analogue acted only as an antagonist of ADP and was without effect on platelet aggregation^{13,14} at physiological pH. Evidence has been presented¹³ which suggests that part of the ineffectiveness of this analogue resided in the pK_a value of the secondary phosphoryl hydroxyl group, which is 8.0 for the analogue compared with 7.0 for the natural nucleotide¹⁵. Additionally, altered stereochemistry of the phosphate chain due to the substitution of a (hydrophobic) methylene group has been invoked to account for the lack of effect of the compound¹⁴.

The K_i value of $\sim 0.3 \mu\text{M}$ for ADP (and ADP- βS) inhibition of adenylate cyclase in the plasma membrane preparation in the present study compares well with the K_i value (0.5 μM) for the inhibition by ADP of cyclic AMP production in intact platelets^{2,6}. However, the amount of inhibition ($\sim 25\%$) elicited by ADP in the present study compares with $\sim 90\%$ inhibition of cyclic AMP production in intact platelets. No immediate explanation for this discrepancy is available, but it may suggest either the unmasking of additional pools of adenylate cyclase or some deregulation of the activity on the disruption of cells and preparation of membranes. There was a similar degree of inhibition of fat cell adenylate cyclase by adenosine analogues which was observed as a dramatic reduction in cellular cyclic AMP levels¹⁶. Sabol and Nirenberg¹⁷ reported similar levels of inhibition of neuroblastoma–glioma hybrid adenylate cyclase by α -adrenergic agents and other opiates.

The effects of a more extensive range of ADP concentrations were also examined. After the sensitive phase of ADP inhibition

Table 1 Effect of ADP and related compounds on platelet adenylate cyclase activity

Additions (2.5 μM)	Cyclic AMP (pmol per min per mg protein)	% of control
—	83.8 (5.5)	100
Adenosine	81.2 (1.6)	97
2'-Deoxy AMP	82.4 (1.76)	98
AP(CH ₂)P	82.0 (1.09)	98
AMP	79.7 (2.56)	95
2'-Deoxyadenosine	77.0 (3.8)	92
2'-Deoxy ADP	75.6 (3.36)	90
ADP	64.6 (3.15)	77
ADP- βS	63.6 (0.63)	76

Adenylate cyclase activity was determined as described in Fig. 1 legend with both 1 μM PGE₁ and 1 μM GTP in the presence of either 2.5 μM ADP or the related compounds indicated. Values presented are the means $\pm$ s.d. of quadruplicate determinations. Activities are also expressed as a percentage of that observed in the absence of the ADP analogues.

which is complete at 2 μM (compare Fig. 1), a plateau in activity follows. Further inhibition is evoked by concentrations $>40 \mu\text{M}$ ($K_i \sim 200 \mu\text{M}$), which presumably represents competition of ADP for substrate at the catalytic site (results not shown).

PGE₁-stimulated activity with and without GTP is shown as a function of time in the presence and absence of 2 μM ADP in Fig. 2. Activities are linear for 10 min in the absence of ADP. Inhibition by ADP is immediate in onset and is maintained throughout the usual assay period (6 min). After 6 min there is some diminution in the amount of inhibition caused by ADP, which may reflect ADP metabolism.

Degradation of ¹⁴C-ADP (2 μM) during the routine assay period (6 min) was assessed by TLC on PEI-cellulose¹² of an aliquot of reaction mixture to which [α -³²P]APP(NH)P had not been added. At least 85% of the radioactivity applied was recovered in the ADP region in two determinations (results not shown). The low level of degradation of ADP by a plasma membrane preparation, presumably containing nucleotide phosphohydrolases, may reflect protection of the diphosphate by the adenylate cyclase substrate, APP(NH)P.

In Table 1 the effect of ADP on platelet adenylate cyclase is compared with that of ADP analogues and related compounds. ADP and ADP- βS yield equal inhibition, AP(CH₂)P is ineffective and 2'-dADP gives intermediate results. The relative ineffectiveness of AMP, dAMP, adenosine and 2'-deoxyadenosine underline the fact that ADP and ADP- βS act on a receptor which is specific for ADP as first suggested by Born¹⁸. The weak inhibition by 2'-deoxyadenosine may be viewed as an interaction with inhibitory adenosine sites described previously^{9,19}.

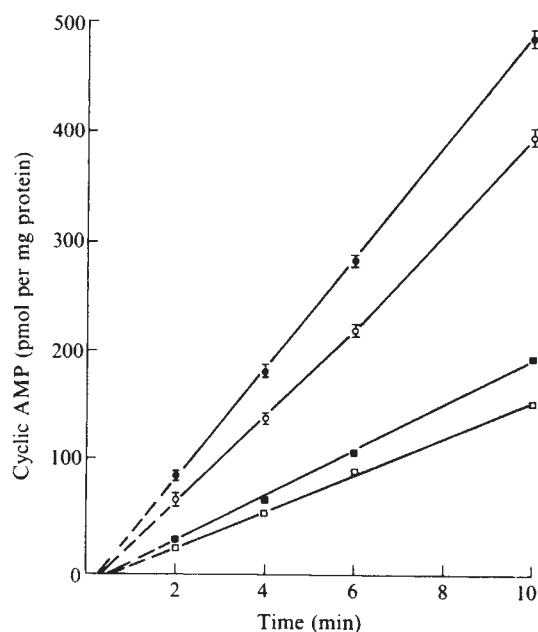


Fig. 2 Time course of the effect of ADP on adenylate cyclase activity of platelet plasma membranes. Cyclic AMP accumulation was measured at the times indicated in the conditions described in Fig. 1 in the presence of 1 μM PGE₁ alone (■, □) or in combination with 1 μM GTP (●, ○) without (●, ■) or with (○, □) 2 μM ADP. Standard deviations of triplicate determinations are shown for the upper two lines.

Table 2 shows the inhibition by ADP of various activity states of platelet adenylate cyclase activity. Clearly the greatest absolute amount of inhibition is obtained with maximally stimulated activities. It is not so clear whether the ADP inhibition is preferentially directed towards a particular state. A notable feature is that the amount of inhibition evoked by ADP is equivalent to that produced by adrenaline in this experiment.

Table 2 Effect of ADP on variously stimulated platelet adenylate cyclase

Additions	Cyclic AMP (pmol per min per mg protein)		% activity in absence of ADP
	-ADP	+ADP (2 μM)	
GTP (1 μM)	4.79 (0.29)	3.5 (0.15)	73
GTP + PGE ₁ (1 μM)	72.9 (4.4)	56.9 (1.02)	78
GTP + EPI (10 μM)	3.74 (0.46)	3.44 (0.33)	92
GTP + EPI + PGE ₁	51.0 (4.1)	44.8 (1.17)	88

Adenylate cyclase activity was determined as described in Fig. 1 in the presence or absence of 2.5 μM ADP with 1 μM GTP, 1 μM PGE₁ and 10 μM (-)-L-adrenaline as indicated. Values presented are the mean $\pm$ s.d. of quadruplicate determinations. Activities are also expressed as a percentage of that observed in the absence of ADP.

Jakobs *et al.*⁸ reported a similar degree of inhibition of platelet adenylate cyclase by adrenaline. Interestingly, the combination of ADP and adrenaline yields a total of 40% inhibition of activity relative to control values (Table 2, 72.9 to 44.8) but the relative amount of inhibition induced by ADP is diminished in the presence of adrenaline and vice versa. These findings suggest that the two agents converge on a common pool of catalytic activity.

The present report has described a direct inhibition of platelet plasma membrane adenylate cyclase activity by ADP in the same concentration range as that which is effective in both reducing cyclic AMP production in intact platelets and promoting platelet aggregation. The detection of this effect of ADP on a membrane adenylate cyclase also opens up the possibility of detecting such effects on the enzyme in other systems. Primary candidates are those physiological systems thought to be mediated by 'purinergic receptors' as described by Burnstock²⁰ and those systems in which adenylate cyclase shows different behaviour when APP(NH)P rather than ATP is used as substrate (compare ref. 21).

ADP inhibits as effectively as adrenaline in this system. A combination of the two agents evokes greater inhibition than that elicited by either agent alone. Precisely how these effectors interact in the regulation of platelet adenylate cyclase remains to be established. However, it is possible that this interaction may serve as a model of the means whereby neurotransmitters and adenine nucleotides, which are commonly secreted together in synapses, participate in the 'neuromodulation' of physiological events²⁰.

We thank Drs G. A. Jamieson and M. Markham for advice and assistance in the preparation of platelet plasma membranes and Drs C. Londres, M. C. Lin and T. Nielsen for helpful criticism.

Received 23 July; accepted 24 September 1979.

1. Cole, B., Robison, A. & Hartmann, R. C. *Ann. N.Y. Acad. Sci.* **185**, 477-487 (1971).
2. Haslam, R. J., Davidson, M. M. L., Davies, T., Lynham, J. A. & McClenaghan, M. D. *Adv. cyclic Nucleotide Res.* **9**, 533-552 (1978).
3. Salzman, E. W. & Levine, L. *J. clin. Invest.* **50**, 131-141 (1971).
4. Mustard, J. F., Packham, M. A., Perry, D. W., Guccione, M. A. & Kinlough-Rathbone, R. L. *Ciba Fdn Symp.* **35**, 47-75 (1975).
5. Kazer-Glanzmann, R., Jakabova, M., George, J. N. & Luscher, E. F. *Biochim. biophys. Acta* **466**, 429-440 (1977).
6. Haslam, R. J. *Ciba Fdn Symposium* **35**, 121-151 (1975).
7. Lee, G. R., Boggs, D. R., Bithell, T. C., Athens, J. W. & Foerster, J. (eds) *Clinical Haematology* 7th edn 372-408. (Kimpton, London, 1974).
8. Jakobs, K. H., Saur, W. & Schultz, G. *FEBS Lett.* **85**, 167-170 (1978).
9. Haslam, R. J., Davidson, M. M. L. & Desjardins, J. V. *Biochem. J.* **176**, 83-95 (1978).
10. Barber, A. J. & Jamieson, G. A. *J. biol. Chem.* **245**, 6357-6365 (1970).
11. Salomon, Y., Londres, C. & Rodbell, M. *Analyt. Biochem.* **58**, 541-548 (1974).
12. Randerath, K. & Randerath, E. *J. Chromatogr.* **16**, 111-129.
13. Horak, H. & Barton, P. G. *Biochim. biophys. Acta* **373**, 471-480 (1974).
14. Gough, G., Maguire, M. H. & Penglis, F. *Molec. Pharmacol.* **8**, 170-177 (1972).
15. Myers, T. C., Nakamura, K. & Danielson, A. B. *J. org. Chem.* **30**, 1517-1520. (1965).
16. Londres, C., Cooper, D. M. F., Schlegel, W. & Rodbell, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5362-5366 (1978).
17. Sabol, S. L. & Nirenberg, M. *J. biol. Chem.* **254**, 1913-1920 (1979).
18. Born, G. V. R. *J. Physiol., Lond.* **209**, 487-511 (1970).
19. Londres, C. & Wolff, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5482-5486 (1977).
20. Burnstock, G. in *Physiological and Regulatory Functions of Adenosine and Adenine Nucleotides* (eds Baer, H. P. & Drummond, G. I.) 3-32 (Raven, New York, 1979).
21. Lin, M. C., Salomon, Y., Rendell, M. & Rodbell, M. *J. biol. Chem.* **250**, 4246-4252 (1975).

Concentration dependence of currents through single sodium-selective pores in frog skin

W. Van Driessche* & B. Lindemann†

* Laboratorium voor Fysiologie, Universiteit Leuven, Belgium

† Abteilung für Membranforschung an Epithelien, Second Department of Physiology, D-6650 Hamburg, FRG

Many transport processes in biological membranes show a saturation of transport rate when the concentration of the transported ions or molecules is sufficiently increased. Saturation will result when at high concentrations a translocation step which does not directly depend on bulk concentration becomes rate limiting, and it is expected for all multi-step translocation mechanisms. However, a possible alternative explanation for saturation is a regulatory process which decreases the number of translocating channels in response to an increase in bulk concentration of the transported species. In the frog epidermis, Na^+ uptake from the extracellular space into the epithelial cytosol occurs through Na^+ -selective, passively conducting channels located in the apical membrane of the cells directly below the cornified outer cell layer. Saturation of transport occurs with increasing outer Na^+ activity ($[\text{Na}]_o$): the net Na^+ current^{1,2}, membrane conductance^{2,3} and unidirectional Na^+ influx⁴ increase only subproportionally with $[\text{Na}]_o$. The apical Na^+ current can transiently be larger than the saturation value, suggesting that a fixed upper single channel transport rate is not responsible for the saturation, and it has, therefore been suggested that in this case saturation occurs through a $[\text{Na}]_o$ -dependent decrease in channel density². The Na^+ -translocation process can be investigated by a statistical evaluation of spontaneous transport fluctuations. The reversible transport blocker amiloride can be used to induce random fluctuations of the Na^+ current at the apical membrane. The power density spectrum of these fluctuations is of the lorentzian type and permits an estimation of the Na^+ current which passes through individual translocators in the time intervals between blocks⁵. We have recorded the current fluctuations at different $[\text{Na}]_o$ values and different amiloride concentrations in the $[\text{Na}]_o$ range in which saturation of total Na^+ current is observed. We report here that the density of the conducting pores decreases with increasing $[\text{Na}]_o$.

The experimental technique was similar to that described previously³. We used abdominal skins of *Rana esculenta* at room temperature in a Lucite chamber with 3 cm² exposed to the bathing solutions. The tissue was held by silicon rings (Silgel 604, Wacker Chemie) to avoid damage to the edges. The inner solution was K^+ -sulphate Ringer, which would be expected to depolarise the K^+ -selective inward-facing membranes of the epithelium and produce a four- to five-fold increase in their

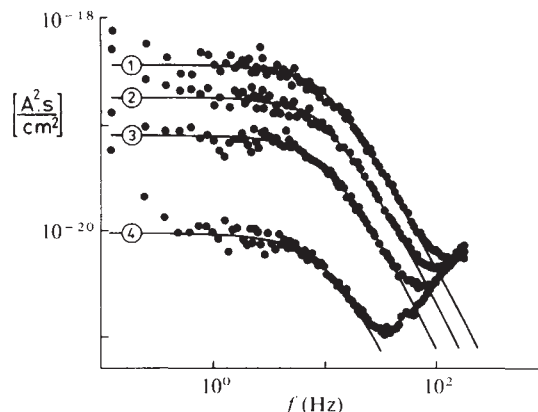


Fig. 1 Power density spectra obtained at constant amiloride concentration (4 μM) and four values of $[\text{Na}]_o$, as indicated. Power densities are normalised to 1 cm². It can be seen that a 10-fold variation in $[\text{Na}]_o$ causes significantly less than a 100-fold variation in S_0^A (plateaus of Lorentz curves), and a small change of the corner frequency f_c^A . In the high frequency range the excess noise becomes smaller than the amplifier noise, which increases because of the decrease in the membrane impedance at high frequencies. Respective values for f_c^A (Hz) and $[\text{Na}]_o$ (mM) are: ① 10.5, 60; ②, 10.4, 30; ③, 9.5, 18; ④, 9.3, 6.

conductance. At high inner K^+ concentrations the epithelium (judged by its current-voltage curve) behaves approximately like a one-membrane preparation². Transepithelial resistance and potential are then largely determined by the Na^+ -selective, outward-facing membrane. This situation was further improved by using amiloride to increase the resistance of this membrane. The outer solution was sulphate Ringer in which the Na^+ concentration was varied by substitution with K^+ . An activity coefficient of 0.55 was used to compute Na^+ activities^{2,6}. We used $[\text{Na}]_o$ values of 6–60 mM and amiloride concentrations, $[\text{A}]_o$, of 1–16 μM for fluctuation analysis. I_{Na} , the transcellular current component which could be blocked by amiloride, was determined as described elsewhere⁵. The transepithelial voltage (V) was clamped to zero with a voltage-clamp circuit designed for low-noise applications⁷. The fluctuating component of the feedback current was amplified and sampled as described previously^{5,8}.

In agreement with previous results, the amiloride-induced component of the spectrum of the Na^+ current was found to have a corner frequency, f_c^A , which increased with the amiloride concentration, whereas the plateau values S_0^A , of the corresponding lorentzian curve decreased. f_c^A increased linearly with $[\text{A}]_o$ values between 3 and 12 μM , and so the amiloride-receptor reaction could be described by first order kinetics using the reaction rate ($1/\tau$)

$$2\pi f_c = \frac{1}{\tau} = k_1'(\text{A})_o + k_2 \quad (1)$$

where k_1' represents the apparent association rate constant, k_2 the dissociation rate constant and τ the relaxation time constant⁵. Corner frequencies smaller than predicted by this straight-line relationship were often found with $[\text{A}]_o$ values of 1, 2 and 16 μM . At these amiloride concentrations the power density spectra usually deviated markedly from the lorentzian, probably because of additive spectral components—at low $[\text{A}]_o$ the amiloride spectrum is too close to the spectrum of the spontaneous fluctuations to be easily evaluated, and at high $[\text{A}]_o$ it is too close to the background spectrum obtained from the epithelium in Na^+ -free solutions. Therefore, we have restricted our analysis to amiloride concentrations between 3 and 12 μM . At constant $[\text{A}]_o$, f_c^A increased by about 1 Hz in response to a 10-fold increase in $[\text{Na}]_o$ (see Fig. 1). With $[\text{Na}]_o = 60$ mM, $2\pi f_c^A$ was calculated to be $62.1 \pm 1.7 \text{ s}^{-1}$ at $[\text{A}]_o = 4 \mu\text{M}$ ($n = 10$), and $115.9 \pm 3.2 \text{ s}^{-1}$ at $[\text{A}]_o = 8 \mu\text{M}$ ($n = 10$). From the shift of f_c^A with

Table 1 The rate constant k_1' and the factor G in equation (2) which are approximately independent of $[\text{Na}]_o$

$[\text{Na}]_o$ (mM)	k_1' ($\mu\text{M}^{-1} \text{ s}^{-1}$)	$G \times 10^3$ (s)	
		$[\text{A}]_o = 4 \mu\text{M}$	$[\text{A}]_o = 8 \mu\text{M}$
6	13.6 ± 0.4	14.8 ± 1.4	7.9 ± 0.5
20	12.4 ± 0.5	14.2 ± 1.2	8.3 ± 0.6
30	13.0 ± 0.6	12.7 ± 0.5	7.9 ± 0.3
45	12.5 ± 0.4	12.0 ± 0.5	6.7 ± 0.2
60	12.2 ± 0.3	13.0 ± 0.7	7.5 ± 0.5

G was calculated for two values of $[\text{A}]_o$. Each value corresponds to the mean $\pm$ s.e.m. of six observations.

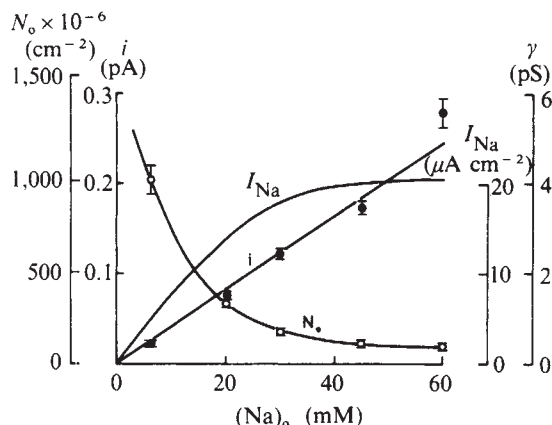


Fig. 2 ●, Computed values (equation (2)) of Na^+ current (i) through single conducting pores at different outer Na^+ activities. Spectra obtained with an amiloride concentration of $6 \mu\text{M}$ were used ($n = 7$). Bars represent s.e.m. The conductance (pS) of single open pores was computed using equation (3). ○, The densities of conducting pores (N_o). The curve labelled I_{Na} represents the total amiloride-sensitive Na^+ current which passed 1 cm^2 of epithelium in the absence of amiloride in the experiment represented by the data points.

$[\text{A}]_o$ (equation (1)) we calculated, for different sodium activities, tentative values for k'_1 , k_2 and the apparent dissociation constant K'_A . k'_1 , which is the slope of the straight line through the data points of the rate/concentration plot, could be determined quite accurately. However, there was a greater experimental variation in k_2 and K'_A , which are, respectively, the intercept with the ordinate and abscissa. Table 1 lists the values for k'_1 obtained at different Na^+ concentrations. k'_1 does not change significantly with $[\text{Na}]_o$, whereas the apparent value of k_2 increases with $[\text{Na}]_o$. The apparent mean lifetime of the amiloride-receptor complex calculated from these k_2 values was $75 \pm 2 \text{ ms}$ and $382 \pm 75 \text{ ms}$, respectively, at $[\text{Na}]_o$ 60 mM ($n = 61$) and 20 mM ($n = 24$).

The current, i , passing through single open pores was calculated from Lorentz spectra using

$$i = S_o^\wedge / (4aGI_{\text{Na}}) \quad (2)$$

where $G = (2\pi f_c^\wedge)^{-2} \times k'_1 \times [\text{A}]_o$. Equation (2) is equivalent to equation (6) derived in ref. 5. a denotes the membrane area and I_{Na} the mean net Na^+ current per cm^2 observed while current fluctuations were being registered. We found that $S_o^\wedge$ and I_{Na} changed with $[\text{Na}]_o$, whereas k'_1 and $f_c^\wedge$ remained almost constant. Therefore, the factor G remains approximately constant as long as $[\text{A}]_o$ does not change (see Table 1). The computed values of i are shown in Fig. 2 as a function of $[\text{Na}]_o$. The current through single conducting pores can be seen to increase linearly with the outer Na^+ activity, although the mean net Na^+ current becomes saturated (Fig. 2, curve labelled I_{Na}). This result supports the previous conclusion² that the transport capacity of individual channels (pores) cannot be responsible for the saturation of net Na^+ uptake with increasing $[\text{Na}]_o$.

The change of i with $[\text{Na}]_o$ enables the conductance of single open pores to be calculated. The current-voltage curve of the Na^+ -selective membrane can be predicted by the constant field equation². Assuming the intracellular Na^+ concentration to be negligible relative to $[\text{Na}]_o$ in the amiloride concentration range used, the slope conductance of the single open pore (γ) can be approximately estimated at zero voltage across the apical membrane as

$$\gamma = \frac{i}{2RT/F} \quad (3)$$

R , T and F have their usual meaning. γ is thus proportional to i , and so increases linearly with $[\text{Na}]_o$. The appropriate ordinate scale for γ has been included in Fig. 2. Mean values of 5.5 pS are reached at $[\text{Na}]_o$ values of 60 mM.

Knowing the dependence of i on $[\text{Na}]_o$, the expected change in the density of conducting pores, N_o , with $[\text{Na}]_o$ can be calculated from the expression $N_o = I_{\text{Na}}/i$. Using this, we do, indeed, find that N_o decreases with increasing $[\text{Na}]_o$ (Fig. 2), thus confirming our hypothesis that the saturation of Na^+ transport with increasing $[\text{Na}]_o$ is caused by a decrease in the number of conducting translocators. This regulatory 'inactivation' of transporting channels may prevent Na^+ overloading of the epithelial cells in situations in which there is a large environmental Na^+ supply.

Received 22 May; accepted 1 October 1979.

1. Cerejido, M., Herrera, F. C., Flanigan, J. W. & Curran, P. F. *J. gen. Physiol.* **47**, 879–893 (1964).
2. Fuchs, W., Hviid Larsen, E. & Lindemann, B. *J. Physiol., Lond.* **267**, 137–165 (1977).
3. Gebhardt, U., Fuchs, W. & Lindemann, B. *Role of Membranes in Secretory Processes*, 284–300 (North-Holland, Amsterdam, 1972).
4. Biber, U. L. T. & Curran, P. F. *J. gen. Physiol.* **56**, 83–99 (1970).
5. Lindemann, B. & Van Driessche, W. *Science* **195**, 292–294 (1977).
6. Owen, H. *The Physical Chemistry of Electrolytic Solutions* (Reinhold, New York, 1967).
7. Van Driessche, W. & Lindemann, B. *Rev. scient. Instrum.* **49**, 32–37 (1978).
8. Van Driessche, W. & Gögelein, H. *Nature* **275**, 665–667 (1978).

Human erythrocyte anion exchange site characterised using a fluorescent probe

James A. Dix, A. S. Verkman & A. K. Solomon

Biophysical Laboratory, Harvard Medical School, Boston, Massachusetts 02115

L. C. Cantley

Department of Biochemistry, and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

A plasma membrane anion exchange system allows the red blood cell to replace intracellular bicarbonate rapidly with extracellular chloride to facilitate carbon dioxide release into the lungs. The protein responsible for this exchange is a glycoprotein of molecular weight about 93,000, called band 3, which comprises approximately 25% of the total red cell membrane protein and is present in the membrane as a non-covalent dimer¹. The molecular mechanism of anion transport remains obscure, in spite of the characterisation of the extracellular anion binding site by a number of methods including interactions with disulphonic stilbene probes which act as specific inhibitors of anion exchange². One of these compounds (4,4'-diisothiocyano-2,2'-disulphonic stilbene; DIDS) causes complete inhibition of anion exchange by reacting at approximately one site per band 3 monomer ($\sim 2.5 \text{ nmol per mg protein}$)³. A fluorescent DIDS analogue, DBDS (4,4'-dibenzamide-2,2'-disulphonic stilbene), binds reversibly and specifically to the DIDS reactive site where it also inhibits anion exchange⁶. When bound, the fluorescence intensity of DBDS increases by two orders of magnitude and this has enabled us to investigate the binding kinetics of DBDS to red blood cell ghost membranes by fluorescence equilibrium and temperature-jump methods. Our equilibrium binding data indicate that there are two classes of DBDS binding sites on red cell ghosts. The reaction scheme that we present here suggests that the apparent high affinity (15–60 nM) for binding the first DBDS molecule is the result of a two step process: an initial binding of lower affinity ($K_1 \sim 0.5 \mu\text{M}$) to band 3 followed by a slow conformational change ($\sim 5 \text{ s}^{-1}$) in the binding site which acts to lock the DBDS molecule in place. A second DBDS molecule binds with low affinity ($\sim 3 \mu\text{M}$) to a site which is more difficult to characterise.

The equilibrium binding affinities for DBDS binding to red cell ghosts were determined as described in Fig. 1 legend. All equilibrium and relaxation experiments were conducted in 28.5 mM sodium citrate, pH 7.4 (160 mM ionic strength). No transportable anions were present so that the system could be

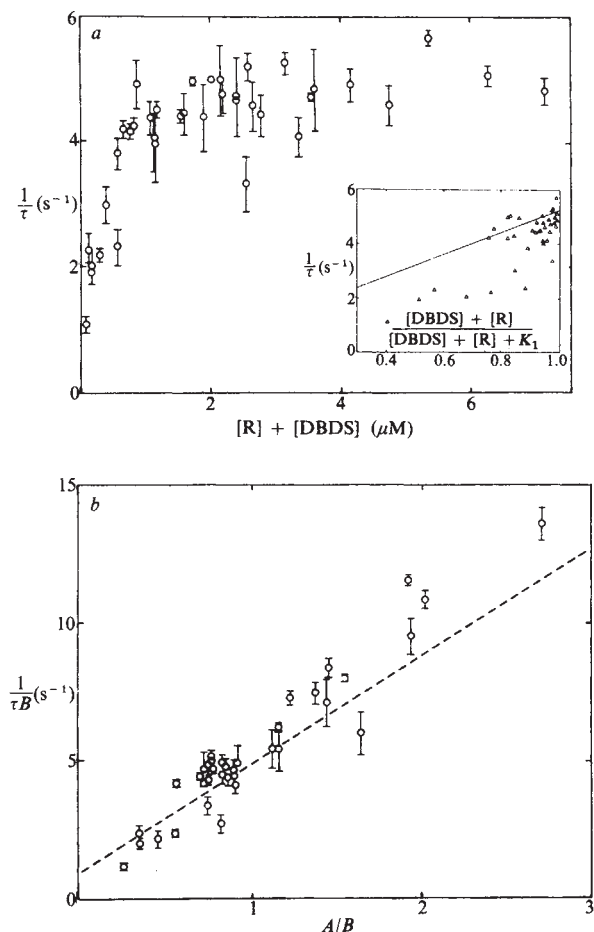


Fig. 3 *a*, Kinetics of DBDS binding to band 3. The abscissa values, the sum of the free concentrations of DBDS and receptor, are determined from the equilibrium binding data shown in Fig. 1 for the independent site model. The ordinate, $1/\tau$, is the reciprocal time constant obtained from a fit to relaxation data. Each data point represents an average of four experiments; error bars indicate one standard deviation from the mean. These data require a minimal mechanism consisting of a rapid bimolecular step followed by a slower conformational change. The relaxation expression for this model (see inset) is $1/\tau = k_{-2} + k_2([\text{DBDS}] + [R])/(K_1 + [\text{DBDS}] + [R])$, where $[\text{DBDS}]$ and $[R]$ are the free concentrations of DBDS and receptor sites calculated from the high affinity independent site model binding data of Fig. 1 and the constants defined in the text. The solid line is the best fit using the assumptions in the text with $k_{-2} = 2.3 \text{ s}^{-1}$ and $k_2 = 3.3 \text{ s}^{-1}$. The effect of the low affinity independent site on the relaxation times predicted by this minimal model have been neglected for simplicity. A fit which includes the effect of the second binding site is similar to that obtained with the minimal model¹⁰. *b*, A fit of the relaxation expression for the three step model to the experimental data. Assuming rapid equilibrium for the two binding steps, K_1 and K_3 , the relaxation rate as determined by standard procedures⁹ is $1/\tau = k_2A + k_{-2}B$ where:

$$A = 1 - \frac{yK_1/2 - [R][\text{DBDS}]}{xy - 2[R]^2[\text{DBDS}]/K_1K_2}$$

$$B = 1 - \frac{x[\text{DBDS}] - [R][\text{DBDS}]/K_2}{xy - 2[R]^2[\text{DBDS}]/K_1K_2}$$

$$x = [R] + [\text{DBDS}] + K_1/2$$

$$y = [\text{DBDS}](1 + 2[R]/K_1K_2) + 2K_3$$

$$K_2 = k_{-2}/k_2$$

The solid line through the data is the best fit ($r = 0.94$) with the constraints that $K_1 = K_1^0(1 + k_2/k_{-2}) = 56 \text{ nM}(1 + k_2/k_{-2})$ and $K_3 = K_3^0/(1 + k_{-2}/k_2) = 1.86 \mu\text{M}/(1 + k_{-2}/k_2)$ as determined from the negative cooperativity model in Fig. 1. $[R]$ and $[\text{DBDS}]$ are the concentrations of unoccupied sites and free DBDS as determined from the data in Fig. 1. The values determined for k_2 and k_{-2} were 4.4 s^{-1} and 1.0 s^{-1} respectively.

locks the first DBDS molecule to the receptor site; subsequently a second DBDS molecule binds to a second site with lower affinity. These studies also illustrate the value of the temperature-jump technique in examining the detailed mechanism of membrane transport reactions.

This work was supported in part by NIH grants GM 15692, GM 26199, and GM 06265 and by a grant from the Research Corporation (L.C.C.). A.S.V. was supported by an Insurance Medical Scientist Scholarship donated by the Prudential Insurance Company of America.

Received 30 July; accepted 24 September 1979.

- Steck, T. L. *J. supramolec. Struct.* **8**, 311–324 (1978).
- Cabantchik, Z. I., Knauf, P. A. & Rothstein, A. *Biochim. biophys. Acta* **515**, 239–302 (1978).
- Ship, S., Shami, Y., Breuer, W. & Rothstein, A. *J. Membrane Biol.* **33**, 311–323 (1977).
- Dodge, J. T., Mitchell, C. & Hanahan, D. T. *Archs Biochem. Biophys.* **100**, 119–130 (1963).
- Lowry, O. H., Rosenbrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- Rao, A., Martin, P., Reithmeier, R. A. F. & Cantley, L. C. *Biochemistry* **18** (in the press).
- Eigen, M. & DeMaeyer, L. in *Techniques of Organic Chemistry* (eds Friess, S. L., Lewis, E. S. & Weissberger, A.) 895–1055 (Wiley, New York, 1963).
- Verkman, A. S., Pandiscio, A. A., Jennings, M. & Solomon, A. K. *Anal. Biochem.* (in the press).
- Czerlinski, G. H. *Chemical Relaxation* (Dekker, New York, 1966).
- Verkman, A. S. thesis, Harvard University, MA (1979).

Distribution of light chains in fast skeletal myosin

Susan Lowey, Pamela A. Benfield, Laura Silberstein & Leslie M. Lang

Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254

Myosin is composed of two heavy chains of molecular weight 200,000 and four light chains of approximate molecular weight 20,000. The heavy chains are organised into a two-headed structure of which each head is completed by the presence of two light chains^{1,2}. In fast skeletal myosin the light chains can be divided into two classes which are chemically and functionally distinct. One member of each class is associated with each head. The class known as the alkali light chain contains two related variants: alkali 1 (A1) and alkali 2 (A2). These light chains are identical in amino acid sequence over their C-terminal 141 residues, but differ in their N-terminal sequences³. The non-integral stoichiometry of light chains in rabbit skeletal myosin^{4,5} first suggested that at least some myosins must be homodimeric (a molecule with the same light chain on each head), but it was not clear to what extent heterodimers (molecules containing both alkali 1 and alkali 2 light chains) also occurred. The most direct approach to determining the distribution of light chains in myosin is to isolate the different myosin isoenzymes. Although ion-exchange chromatography has been successful in separating the single-headed sub-fragments of myosin⁶, it has not been possible to fractionate double-headed species except by some type of affinity chromatography^{7,8}. We have shown that homodimers of myosin can be isolated from chicken pectoralis myosin by immunological methods^{9,10}. By comparing the electrophoretic patterns of homodimeric molecules with whole myosin in non-denaturing gels, it has been possible to demonstrate conclusively the existence of a light chain heterodimer.

Antibodies specific for the N-terminal sequence of alkali 1 or alkali 2 can be used as immunoadsorbents to fractionate heavy meromyosin or myosin into two components^{9,10}. When myosin is passed through an affinity column prepared by coupling antibody to Sepharose, a fraction rich in alkali 1 (Fig. 1*b*) or alkali 2 (Fig. 1*c*) is obtained in the void volume, depending on the specificity of the column. The retained myosin fraction can only be eluted by a denaturant such as guanidine hydrochloride, or a weak acid. It is evident that the isolated homodimers are both slightly contaminated by the heterologous light chain: that is, A1 myosin contains about 10% alkali 2 light chain (Fig. 1*b*) and A2 myosin is contaminated by a similar amount of alkali 1 light chain (Fig. 1*c*). The antibody-bound myosin (not shown

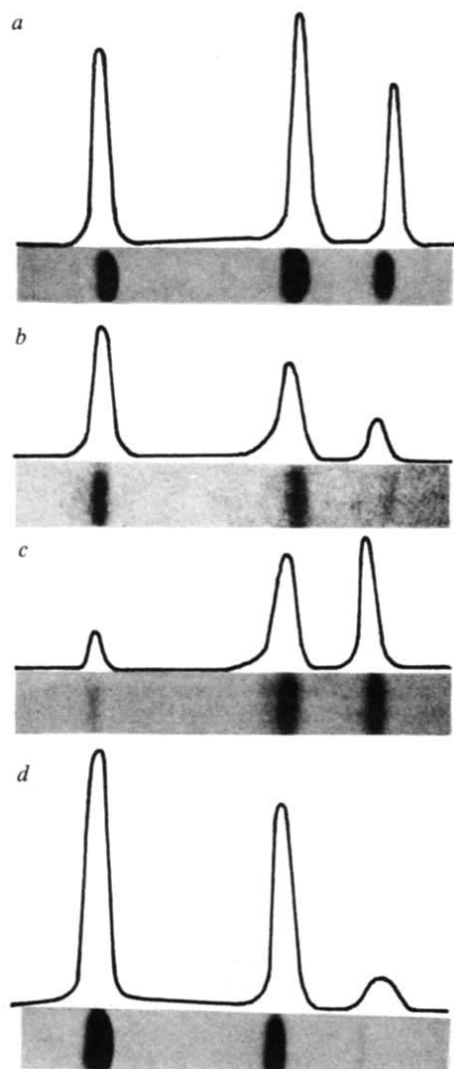


Fig. 1 SDS polyacrylamide gel electrophoresis of chicken pectoralis myosin. A separating gel 12.5% in acrylamide (pH 8.8) and a stacking gel 5% in acrylamide (pH 6.8) was used according to the method of Laemmli²⁴. Gels were stained in 0.025% Coomassie brilliant blue (R250) dissolved in 45% methanol and 9% acetic acid, and destained by diffusion in 5% methanol and 7.5% acetic acid. Myosin loads are about 10 μ g. Electrophoresis is from left to right (positive electrode) separating light chains into A1, DTNB and A2, respectively. *a*, Myosin control; *b*, A1 myosin; and *c*, A2 myosin were prepared from chicken pectoralis myosin by using immunoadsorbents specific for A2 light chain and A1 light chain, respectively (see refs 9 and 10); *d*, A1-exchanged myosin was prepared from chicken pectoralis myosin by addition of a tenfold molar excess of A1 light chain in the presence of 4.7 M NH_4Cl , as described by Wagner and Weeds¹⁵. Excess light chain was removed by gel filtration on Sephacryl S-200 (Pharmacia). Gels were traced on a Joyce Lobel double-beam recording microdensitometer.

here, see refs 9 and 10), is contaminated by an even greater amount of heterologous light chain. If we assume that all the heterologous light chain in both the retained and unretained fractions is derived from a heterodimeric species, a maximum estimate of about 30% A1, A2 myosin is calculated¹¹. However, any nonspecific binding of myosin to these immunoadsorbents would result in an overestimate of this figure. To establish with any confidence that a significant proportion of the myosin population consists of heterodimer, it was necessary to apply an independent approach to this problem.

The studies of Hoh *et al.* have provided such an alternative method by demonstrating that myosin solubilised in alkaline pyrophosphate buffer can be fractionated in low percentage polyacrylamide gels in the complete absence of denaturing agents¹². In such a system, where myosin migrates according to its intrinsic mobility at that pH, chicken pectoralis myosin was

found to separate into three bands. The material from each band was isolated by slicing the gel, and the light chain pattern was analysed by SDS polyacrylamide gel electrophoresis. It was found that the least mobile band contained only alkali 1 light chain, the most mobile band had only alkali 2 light chain, and the middle band contained both light chains. These results suggested that the band of intermediate mobility might be a light chain heterodimer^{13,14}. However, it was not clear to what extent heavy chain heterogeneity might contribute to the band separation in this system, and it was possible that the band of intermediate mobility might arise from a mixture of light chain homodimers with different heavy chains. Our ability to prepare relatively pure homodimers of myosin allowed us to investigate directly whether the heavy chains contribute to the mobility of the isoenzymic species.

Figure 2 shows non-denaturing gels of native chicken pectoralis myosin and the homodimeric myosin molecules isolated by immunoadsorption as described above. The unfractionated myosin shows three bands (Fig. 2*a*), whereas A1 myosin (Fig. 2*b*) and A2 myosin (Fig. 2*c*) each migrate as a single band. Moreover, the A1 isoenzyme co-migrates with the slowest moving component in the control (Fig. 3*b*) and the A2 isoenzyme co-migrates with the fastest species (Fig. 3*c*). To demonstrate that the intermediate band was not generated by light chain exchange in the slightly alkaline conditions of electrophoresis, A1 and A2 myosin were mixed and applied to one gel. No intermediate band was observed (Fig. 2*d*). Similarly, no intermediate band was generated when the two isoenzymes were dialysed together for 24 h against pyrophosphate electrophoresis buffer. It thus seems reasonable to conclude that the band of intermediate mobility represents a light chain heterodimer, and is not simply the result of heavy chain heterogeneity or an experimental artefact.

These experiments do not rule out the possibility that the heavy chains contribute to the mobility of the individual myosin isoenzymes in this gel system. If, however, myosin mobility is solely a function of the alkali light chain, it should be possible to generate a single electrophoretic species from native pectoralis myosin by exchanging the native light chains with either alkali 1 or alkali 2. Such an exchange has been shown to occur in myosin subfragment-1 when it is incubated with an excess of alkali light

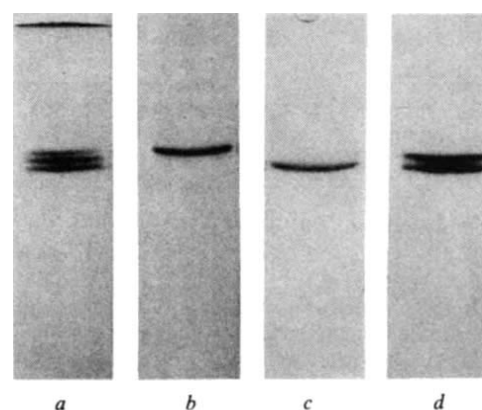


Fig. 2 Native myosin on non-denaturing gels. Polyacrylamide gels (4% acrylamide, 3% bisacrylamide) were prepared in the absence of denaturing agents in 40 mM pyrophosphate buffer (pH 8.7), 1 mM EDTA, 2 mM cysteine, 10% glycerol according to the method of Hoh¹². Electrophoresis was carried out at a constant current of 6 mA per gel for 16 h with buffer recirculating at 60 ml per min. Gels were stained in 0.04% Coomassie brilliant blue (G250) dissolved in 3.5% perchloric acid²⁵, and destained by diffusion in 5% methanol and 7.5% acetic acid. Myosin loads range from 1 to 3 μ g. Positive electrode was at the bottom of the gel. Samples were prepared as described in Fig. 1. *a*, Myosin control is resolved into A1 homodimer, A1,A2 heterodimer, and A2 homodimer (see text); *b*, A1 homodimer, and *c*, A2 homodimer each electrophorese as a single band, while the mixture (*d*) of both isoenzymes resolves into a two-band pattern lacking the middle band found in control myosin.

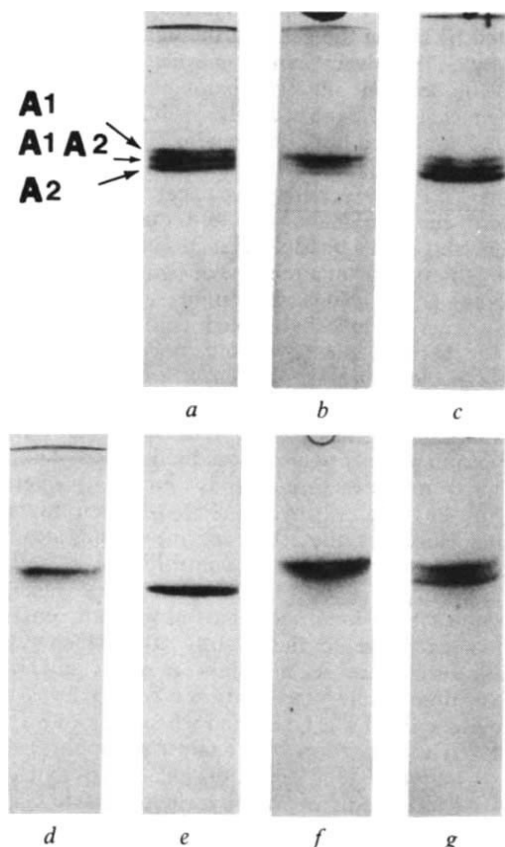


Fig. 3 Co-migrations and light chain-exchanged myosins on non-denaturing gels. Gels were run as previously described in Fig. 2. *a*, Control myosin shows three bands of approximately equal intensity; *b*, impure sample of myosin eluted in the void volume from an antibody affinity column specific for A2 light chain. There is an enrichment of the band of slowest mobility indicating that this contains A1 homodimer. *c*, Co-migration of control myosin and A2 myosin isolated from an antibody affinity column specific for A1 light chain. The enrichment of the band of fastest mobility shows that this represents A2 homodimer. Exchange of total (control) myosin with a tenfold molar excess of A1 light chain in 4.7 M NH_4Cl resulted in myosin of uniform mobility (*d*) which co-migrated with the single band from A1 myosin. A similar exchange was carried out using A2 myosin as the starting material (*e*) and adding an excess of A1 light chain in 4.7 M NH_4Cl . The exchanged myosin again ran as a single band (*f*) with a shift in mobility from that of the A2 myosin. Co-migration of A1-exchanged myosin and A2 myosin (*g*) resulted in two bands.

chain in the presence of 4.7 M NH_4Cl (ref. 15). By this means, a myosin species was generated from total chicken pectoralis myosin which contained only the alkali 1 light chain as determined by SDS gel electrophoresis (Fig. 1*d*). Electrophoresis of this myosin in the non-denaturing gel system shows primarily a slow-moving A1 myosin component, and trace amounts of the band of intermediate mobility (Fig. 3*d*). A similar result was obtained by exchanging A2 myosin homodimer produced by antibody affinity chromatography (Fig. 3*e*) with A1 light chain. A myosin species was generated whose mobility (Fig. 3*f*) was indistinguishable from that of naturally occurring A1 myosin homodimer.

The results presented here strongly suggest that in chicken pectoralis myosin the alkali light chain content is the principal determinant of mobility differences on non-denaturing polyacrylamide gels in pyrophosphate buffer. Therefore, the band of intermediate mobility must be an alkali light chain heterodimer, as suggested^{13,14}. Chicken pectoralis myosin contains approximately equal amounts of alkali 1 and alkali 2 light chains^{9,11}. If heavy chains associated randomly with light chains, one would predict that 50% of the myosin from this muscle would be heterodimeric for the light chains, 25% homodimeric for alkali 1 and 25% homodimeric for alkali 2. It is difficult to quantify the

relative amounts of each component from densitometer traces of the gels shown in Fig. 2, as the resolution is poor, but we would estimate that there are approximately equal amounts of each species. A similar estimate has been made from pyrophosphate gel densitometry by Hoh¹³ and d'Albis¹⁴, and from quantification of material eluted from antibody affinity columns¹¹. These values suggest, therefore, that light chain-heavy chain association is not strictly random.

Heavy chain heterogeneity has been observed in certain peptides isolated from rabbit myosin^{16,17}; however, both of the heavy chain *N*-terminal peptides were found to be associated with each of the two subfragment-1 isoenzymes¹⁸. These findings, together with the observation that the actin-activated ATPase activity appears to depend on the type of alkali light chain present¹⁵, argue against a specific interaction between the light and heavy chains of myosin. Nevertheless, these data cannot exclude the possibility that the alkali light chain homodimers may yet prove to differ from each other in heavy chain sequence, in spite of the absence of a contribution by the heavy chain to the electrophoretic mobility.

Precedents exist for both heavy chain homodimers and heterodimers. In chicken anterior latissimus dorsi myosin¹³, and rat ventricular myosin¹⁹, multiple band patterns on pyrophosphate gels have been demonstrated, in spite of the existence of only one light chain species analogous to the alkali light chains. In the case of rat ventricular myosin, peptide mapping of three isolated species showed the presence of two types of heavy chain homodimers as well as a heterodimeric myosin¹⁹.

Homodimers of heavy chains have also been shown to occur in the nematode worm *Caenorhabditis elegans*²⁰. Moreover, the synthesis of the two types of heavy chains is coordinated throughout larval development, and reflects the final accumulation of two populations of homodimeric molecules²¹. No heterodimers are detectable in the nematode. In contrast, the alkali 1 and alkali 2 light chains are not simultaneously synthesised during early development of vertebrate skeletal muscles and alkali 2 is largely absent in the prenatal stages of the rat²² and chicken²³. It is possible that uncoordinated synthesis of the alkali light chains persists in adult muscle cells, and this may account for the non-random distribution of light chains among the isoenzyme species. Alternatively, the affinity of the light chains for the heavy chains may be such that the formation of a heterodimer is less favoured than the more symmetrical molecules.

Although no single satisfactory explanation can be offered for the occurrence of light chain homodimers and heterodimers in fast skeletal myosin, the proof of their existence at least raises the intriguing question of their function in muscle contraction.

This study was supported by grants AM-17350 from the USPHS, PCM75-4790 from the NSF and a grant (S.L.) and fellowship (P.A.B.) from the Muscular Dystrophy Association.

Received 27 June; accepted 26 September 1979.

- Lowey, S. & Risby, D. *Nature* **234**, 81–85 (1971).
- Weeds, A. G. & Lowey, S. *J. molec. Biol.* **61**, 701–725 (1971).
- Frank, G. & Weeds, A. G. *Eur. J. Biochem.* **44**, 317–334 (1974).
- Weeds, A. G., Hall, R. & Spurway, N. C. *FEBS Lett.* **49**, 320–324 (1975).
- Sarkar, S. *Cold Spring Harb. Symp. quant. Biol.* **37**, 14–17 (1972).
- Weeds, A. G. & Taylor, R. S. *Nature* **257**, 54–56 (1975).
- Wagner, P. D. *FEBS Lett.* **81**, 81–85 (1977).
- Trayer, H. R., Winstanley, M. A. & Trayer, I. P. *FEBS Lett.* **83**, 141–144 (1977).
- Holt, J. C. & Lowey, S. *Biochemistry* **16**, 4398–4402 (1977).
- Lowey, S., Silberstein, L., Gauthier, G. F. & Holt, J. C. in *Motility in Cell Function* (Academic, New York, in the press).
- Silberstein, L. thesis, Brandeis Univ. (1979).
- Hoh, J. F. Y., McGrath, P. A. & White, R. I. *Biochem. J.* **157**, 87–95 (1976).
- Hoh, J. F. Y. *FEBS Lett.* **90**, 297–300 (1978).
- d'Albis, A., Pantaloni, C. & Béchet, J. J. *Eur. J. Biochem.* **99**, 271–272 (1979).
- Wagner, P. D. & Weeds, A. G. *J. molec. Biol.* **109**, 455–473 (1977).
- Weeds, A. G. *Biochem. J.* **104**, 44P (1967).
- Starr, R. & Offer, G. W. *J. molec. Biol.* **81**, 17–31 (1973).
- Pope, B. J., Wagner, P. D. & Weeds, A. G. *J. molec. Biol.* **109**, 470–473 (1977).
- Hoh, J. F. Y. & Yeoh, G. P. S. in *Fibrous Proteins: Scientific, Industrial and Medical Aspects* (Academic, New York, in the press).
- Schachat, F., Garcea, R. L. & Epstein, H. F. *Cell* **15**, 405–411 (1978).
- Garcea, R. L., Schachat, F. & Epstein, H. F. *Cell* **15**, 421–428 (1978).
- Gauthier, G. F., Lowey, S. & Hobbs, A. W. *Nature* **274**, 25–29 (1978).
- Chi, J. C. H., Rubinstein, N., Strahs, K. & Holtzer, H. *J. Cell Biol.* **67**, 523–537 (1975).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Reisner, A. H., Nemes, P. & Bucholtz, C. *Analyt. Biochem.* **64**, 509–516 (1975).

Nucleotide sequence of a cDNA clone encoding human preproinsulin

Graeme I. Bell, William F. Swain, Raymond Pictet, Barbara Cordell, Howard M. Goodman & William J. Rutter

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

Insulin consists of two polypeptide chains, A (21 amino acids) and B (30 amino acids), linked by disulphide bonds. Both chains are derived from one precursor, proinsulin, which includes a connecting peptide (C) between the A and B chains, and which is excised before the secretion of insulin from the pancreatic B cells¹. The observation that the C-peptide varies between species, in contrast to the highly conserved A and B sequences¹⁻³, is consistent with the theory that it serves a purely structural function in insulin synthesis. During *in vitro* translation of insulin mRNA, a larger peptide containing about 25 additional residues at the N-terminal end (preproinsulin) is the primary product⁴⁻⁷. The prepeptide is cleaved to leave proinsulin during transport into the endoplasmic reticulum and is thought to direct this process specifically⁸. Using automated amino acid sequence analysis, the partial amino acid sequences of the prepeptide regions of bovine, rat, sea raven and anglerfish preproinsulin have been established^{4,5,7}. The nucleotide sequences of the cloned cDNA and gene coding for rat insulin I have confirmed the amino acid sequence of rat proinsulin I, and have also predicted the sequence of the prepeptide⁹⁻¹¹. Like the prepeptides of other secreted proteins, this prepeptide has a prominent hydrophobic region¹². We report here the cloning of a cDNA prepared from human insulin mRNA and an analysis of the nucleotide sequence of the cloned molecule including the region coding for the prepeptide and portions of the 5'- and the 3'-untranslated regions of the molecule. We also compare the structure of the human molecule with the previously reported rat mRNA⁹⁻¹¹.

The mRNA used in these experiments was prepared from a human insulinoma by the guanidinium thiocyanate procedure⁹. Approximately 40 ng of double-stranded cDNA was prepared from 130 µg of unfractionated RNA and inserted into the *Pst*I endonuclease site of the β -lactamase gene of pBR322 (ref. 13) using poly(dG)-poly(dC) homopolymeric extensions. The *Escherichia coli* strain X1776 was transformed with this DNA in P2 containment conditions according to the NIH Guidelines. This produced 525 tetracycline-resistant transformants. These were replica-plated and screened for insulin sequences by a modification of the colony hybridisation method of Grunstein and Hogness¹⁴ using nick translated cloned rat preproinsulin I cDNA^{9,11} as a probe. Since the amino acid sequence of rat and human insulins are quite similar (insulin, 92% homology and

```

-24      -20
met ala leu trp met arg leu leu pro leu leu ala leu
UCCUUCUGCC AUG GCC CUG UGG AUG CGC CUC CUG CCC CUG CUG GCG CUG

-10      1
leu ala leu trp gly pro asp pro ala ala ala phe val asn gln his
CUG GCC CUC UGG GGA CCU GAC CCA GCC GCA GCC UUU GUG AAC CAA CAC

10      20
leu cys gly ser his leu val glu ala leu tyr leu val cys gly glu
CUG UGC GGC UCA CAC CUG GUG GAA GCU CUC UAC CUA GUG UGC GGG GAA

30
arg gly phe phe tyr thr pro lys thr arg arg glu ala glu asp leu
CGA GGC UUC UAC ACA CCC AAG ACC CGC CGG GAG GCA GAG GAC CUG

40      50
gln val gly gln val glu leu gly gly gly pro gly ala gly ser leu
CAG GUG GGG CAG GUG GAG CUG GGC GGC GGC CCU GGU GCA GGC AGC CUG

60
gln pro leu ala leu glu gly ser leu gln lys arg gly ile val glu
CAG CCC UUG GCC CUG GAG GGC UCC CUG CAG AAG CGU GGC AUU GUG GAA

70      80
gln cys cys thr ser ile cys ser leu tyr gln leu glu asn tyr cys
CAA UGC UGU ACC AGC AUC UGC UCC CUC UAC CAG CUG GAG AAC UAC UGC

86
asn AM
AUC UAG ACGCAGCCCGCAGGCAGCCCGCCCGCCGCCUCCUGCACCAGAGAGAUGGA

AUAAAGCCCUUGAACACG

```

Fig. 2 Sequence of human preproinsulin mRNA. The sequence of the mRNA was deduced from that of the cDNA insert in pcHI-1. The predicted amino acid sequence is indicated above. The prepeptide region extends from the Met at position -24 to the Ala at position -1, the B-peptide from the Phe at position 1 to the Thr at position 30, the C-peptide from the Arg at position 31 to the Arg at position 65, and the A peptide from the Gly at position 66 to the Asn at 86. This cDNA insert also contains a poly(A) tract of approximately 45 bases extending from the 3'-terminal cytosine.

proinsulin, 83% homology), it was anticipated that the cloned rat cDNA should cross-hybridise with human insulin sequences in conditions of reduced stringency. After hybridisation as described¹¹, the filters were washed in a solution of 0.015 M sodium chloride, 0.0015 M sodium citrate, 0.1% SDS at 42 °C for 1 h. Autoradiography revealed that 2 out of the 525 colonies hybridised with the cloned rat preproinsulin I probe. Both colonies were ampicillin sensitive and contained plasmids that were approximately 250 and 500 base pairs larger than pBR322, the plasmid vector. The larger plasmid, pcHI-1, was analysed further.

The nucleotide sequence of the inserted cDNA fragment was determined by the procedure of Maxam and Gilbert¹⁵. The overall structure of the cDNA clone pcHI-1 and the sequencing strategy used are depicted in Fig. 1. The nucleotide sequence of both DNA strands was determined as indicated and the sequence of the remaining portions were done in duplicate; in all cases sufficient data were obtained to overlap the junctions of the fragments. The mRNA sequence derived from the cDNA

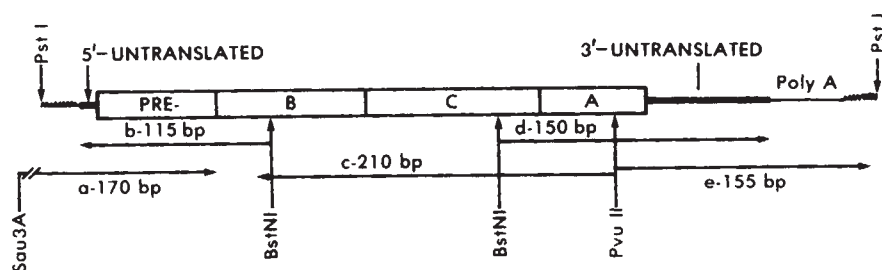


Fig. 1 Organisation and sequence strategy for the human preproinsulin cDNA insert in pcHI-1. The regions coding for preproinsulin are indicated by the open boxes. The wavy lines indicate the poly(dG)-poly(dC) tails of the cDNA insert (27 base pairs at the 5' end and 16 base pairs at the 3' end of the cDNA insert, respectively). Only the restriction sites used for sequencing are indicated. The 5'-end-labelled restriction fragments were derived from the intact plasmid and after the appropriate secondary endonuclease digestion

used for sequencing by the procedure of Maxam and Gilbert¹⁵ using five reactions: G, G+A, A+C, T+C, C. The number of bases determined from each sequencing experiment starting at the indicated restriction site is shown above the arrow. The *Sau*3A site is in the β -lactamase gene 58 base pairs from the *Pst*I site¹³. The size of human preproinsulin mRNA is approximately 550 base pairs as determined by hybridisation after electrophoresis in CH₃HgOH-agarose gels¹⁶. The cloned cDNA (cHI-1) is only 461 base pairs (including 45 A residues) and thus does not represent a complete copy of the mRNA.

Table 1 Comparison of human and rat I and II preproinsulins and their mRNAs

Region	Homology (relative to human)			
	Amino Acid		Nucleotide	
	Rat I	Rat II	Rat I	Rat II
5'-Untranslated	—	—	5/10 (50%)	—
Pre-peptide	19/24 (79%)	—	58/72 (81%)	—
B-Chain	27/30 (90%)	27/30 (90%)	75/90 (83%)	—
C-Peptide	25/35 (71%)	26/35 (74%)	76/105 (72%)	—
A-Chain	20/21 (95%)	20/21 (95%)	59/63 (94%)	57/63 (90%)
3'-Untranslated	—	—	See text	

binding¹⁷; nevertheless there are nucleotide changes in these regions (Fig. 3).

Protein sequence studies indicate that the C-peptide is the most variable portion of proinsulin¹⁻³. Although the length of the C-peptide varies in other species¹⁻³, it is the same in human and rat (35 amino acids; 71% homology). The nucleotide sequence homology is 72%. The prepeptide portions of human insulin and rat insulin I are more homologous than the C-peptides (79% amino acid; 81% nucleotide). Most of the differences that occur in the prepeptides are in the carboxyl terminal portion. Indeed the first 52 bases near the amino terminus seem to be as highly conserved as the B and A chains (48/52 bases = 96%), whereas the latter 20 bases are much less conserved (10/20 bases = 50%). This suggests that this portion of the prepeptide has a function. Examination of the first 24 amino acid residues of the molecules by the sequence-predictive method of Chou and Fasman²¹ shows both α -helix and β -forming potential but the former is stronger; thus two helical segments (residues -24 to -17 and -16 to -8), connected by the helix-disrupting Pro (-16), are suggested. These are the regions with the highest degree of sequence conservation. A comparison of prepeptides of other secretory proteins shows a considerable variation in the amino acid sequences, however there is always a hydrophobic region in this type of structure²². The strong amino acid and nucleotide sequence conservation in this region of preproinsulin would suggest some element of specificity associated with this region of the molecule. It seems unlikely that the conserved structure is concerned with the cleavage reaction since preparations of signal peptidase²³ are able to cleave molecules with different prepeptide sequences. The relatively large area of homology in the prepeptide and the possible secondary structure are more suggestive of interaction with a receptor site. Receptor-mediated endocytosis has been implicated in the specific internalisation of a number of proteins from outside the cell (see ref. 24 for refs). A receptor-mediated mechanism may also be operative at the site of internalisation of the peptides in the endoplasmic reticulum. The lack of specificity seen in the sequestration of a number of secretory proteins in pancreatic microsomes^{7,25,26} could be due to a reduced degree of specificity of the receptors in this system. Alternatively, the true specificity may not have been apparent because the rates of internalisation of proteins were not measured in these experiments.

Outside the region of the mRNA encoding the preproinsulin peptide the sequences of human and rat insulin I diverge rapidly. There is only 50% homology in the short region of the 5'-untranslated sequence in the clone. The clone probably lacks about 50 bases at the 5' end that are present in the mature mRNA. The 3'-untranslated region of human preproinsulin cDNA is 21 residues longer than the cognate region of rat preproinsulin I. Despite the difference in length and the sequence divergence of this region, there is a 21-nucleotide segment adjacent to the poly(A) attachment site that is highly conserved (18/21 = 86%). These 21 nucleotides include the AAUAAA sequence found in many eukaryotic mRNAs²⁷. A similar conservation of a segment of the 3'-untranslated region including the AAUAAA sequence has been observed in the mRNAs of human and rabbit β -globin²⁸ and in the human and rat growth hormones¹⁸⁻²⁰, although the sequences conserved are different in each case. This suggests another gene-specific ele-

ment of regulation which could operate at the level of transcription, translation or polyadenylation.

The cloned human insulin cDNA will be helpful in the analysis of human genomic DNA segments which encode insulin. Moreover, it should be useful for programming the expression of human insulin in alternative hosts, such as bacteria or yeast.

We thank Dr John Karam for providing the human insulinoma, Drs W. Wu and J. T. Yang for helpful discussions, Leslie Spector for assistance in preparing the manuscript, and Jeffrey Edman and Edmund Tischer for technical assistance. This work was supported by a grant from Eli Lilly Co. H.M.G. is an Investigator of the Howard Hughes Medical Institute.

Received 19 July; accepted 17 September 1979.

1. Tager, H. S. & Steiner, D. F. *A. Rev. Biochem.* **43**, 509-538 (1974).
2. Nicol, D. S. H. W. & Smith, L. F. *Nature* **181**, 483-485 (1960).
3. Oyer, P. E., Cho, S., Peterson, J. D. & Steiner, D. F. *J. biol. Chem.* **246**, 1375-1386 (1971).
4. Chan, S. J., Keim, P. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1964-1968 (1976).
5. Lomedico, P. T., Chan, S. J., Steiner, D. F. & Saunders, G. F. *J. biol. Chem.* **252**, 7971-7978 (1977).
6. Permutt, M. A. & Routman, A. *Biochem. biophys. Res. Commun.* **78**, 855-862 (1977).
7. Shields, D. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2059-2063 (1977).
8. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835-861 (1975).
9. Ullrich, A. *et al. Science* **196**, 1313-1319 (1977).
10. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727-3731 (1978).
11. Cordell, B. *et al. Cell* **18**, 533-544 (1979).
12. Devillers-Thiery, A., Kindt, T., Scheele, G. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5016-5020 (1975).
13. Sutcliffe, J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3737-3741 (1978).
14. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965.
15. Maxam, A. & Gilbert, W. *Meth. Enzym.* (in the press).
16. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
17. DeMeyts, P., Van Obberghen, E., Roth, J., Wollmer, A., & Brandenburg, D. *Nature* **273**, 504-509 (1978).
18. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 486-494 (1977).
19. Martial, J. A., Hallewell, R. A., Baxter, J. D. & Goodman, H. M. *Science* **205**, 602-607 (1979).
20. Goodman, H. M., Seeburg, P. H., Shine, J., Martial, J. A. & Baxter, J. D. *Alfred Benzon Symp.* **13**, Munksgaard (in the press).
21. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211-222, 222-245 (1974).
22. Chan, S. J. *et al. From Gene to Protein: Information Transfer in Normal and Abnormal Cells* (Academic, New York, in the press).
23. Jackson, R. C. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5598-5602 (1977).
24. Goldstein, J. L., Anderson, R. G. W. & Brown, M. S. *Nature* **279**, 679-685 (1979).
25. Vishwanath, R. L., Devillers-Thiery, A. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2432-2436 (1977).
26. Shields, D. & Blobel, G. *J. biol. Chem.* **253**, 3753-3756 (1978).
27. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).
28. Kafatos, F., Efstratiadis, A., Forget, B. & Weissman, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5618-5622 (1977).

N-terminal amino acid sequencing of prolamins from wheat and related species

Jean-Claude Autran*, Ellen J.-L. Lew, Charles C. Nimmo & Donald D. Kasarda

Food Proteins Research Unit, Western Regional Research Center, Agricultural Research, SEA, US Department of Agriculture, Berkeley, California 94710

The gliadins of common bread wheat (*Triticum aestivum* L.) make up the major storage protein fraction of grain endosperm and comprise a complex mixture of proteins with similar amino acid compositions and properties¹. Two-dimensional methods of electrophoresis^{2,3}, in which separations are based mainly on net charge, separate gliadins of a single wheat variety into 30-46 components and there is considerable variation among the electrophoretic patterns of different wheat varieties¹. The gliadins seem to be encoded by clusters of duplicated genes² and may constitute a multigene family⁴. The fact that several purified gliadin components have closely similar N-terminal amino acid sequences^{5,6} led us to consider the possibility that, despite the complexity of the mixture, gliadins have sufficient homology in their N-terminal sequences to allow estimation of the homology by sequencing the unfractionated mixture. Here, we provide

* Permanent address: Institut National de la Recherche Agronomique, Laboratoire de Technologie des Blés Durs et du Riz, 34060 Montpellier, France.

evidence for two major groups of gliadins in wheat based on N-terminal sequencing, show that similar groups of prolamins (a general term for equivalent proteins in other species) are found in related species of *Triticum* and *Aegilops* that may have contributed genomes to polyploid wheats with *ABD* genome composition, and show that only one of these groups is found in rye (*Secale cereale* L.).

Wheat and other grain samples used are listed in Table 1. The prolamins of common wheat, *Triticum monococcum*, and rye were prepared from endosperm only (flour) whereas those of other samples, of which we had only small amounts, were prepared from de-germed grain that was ground in a Wiley mill. Prolamins were prepared according to the procedure of Charbonnier⁷ except that the final product was dissolved in 0.01 M acetic acid or 60% ethanol-water and dialysed thoroughly against distilled water. The small amount of precipitate that formed (5% or less of the amount dissolved) was removed by centrifugation and the clear supernatant was freeze-dried. Rye prolamins were less soluble in water—about two-thirds of the preparation precipitated—and both the precipitate and the supernatant were freeze-dried separately for use in sequencing. Yields of prolamins ranged from 3 to 8% of starting material. Electrophoretic patterns of typical preparations, which illustrate the complexity of the mixtures, are shown in Fig. 1.

Automatic sequencing⁸ was carried out on a Beckman sequencer model 890B with DMAA peptide program 111374. Eight cycles were considered sufficient for our purposes. (One sample, *T. monococcum*, was sequenced only through six cycles.) Residue identification was by gas chromatography (GC) according to the procedure of Pisano *et al.*⁹. In all cases, only a few phenylthiohydantoin (PTH) amino acids (sometimes only one) were identified at each cycle of the Edman degradation and recoveries were sufficient to account for 70–80% of the protein sample when moisture contents (10%), probable sequencer yields (65% assumed¹⁰) and prolamins molecular weights (36,000 assumed¹) were taken into account. Typical results are given in Table 2.

The major PTH amino acids at each cycle usually corresponded to the sequence Val-(Arg?)–Val-Pro-Val-Pro-Gln-Leu- for all species of *Triticum* and *Aegilops* except for the wheat variety 'Justin' and an accession of *Triticum urartu* (G-1876), where it was the minor sequence. This sequence corresponds to that of α -gliadins^{5,6}; we shall refer to it as the α -type sequence, although Bietz *et al.*⁶ found this sequence also in a β -gliadin and a γ -gliadin (γ_1 -gliadin). Arg and His were not determined by our GC method. We assume Arg to be an important amino acid at position 2 by analogy with the α -gliadin sequence^{5,6} and identified it at cycle 2 in two samples by liquid chromatography.

A second sequence corresponding to Asn-(Met/Ile)-Gln-Val-(Val/Asp)-Pro-Gln-Gly- was present in all the samples of *Triticum* and *Aegilops*. This sequence is similar to the γ_2 - and γ_3 -gliadin sequences described by Bietz *et al.*⁶ (Table 2); we shall refer to it as the γ -type sequence. It was the minor sequence in most samples except for 'Justin' and *T. urartu* G-1876, where it was present in amounts about equal to (or

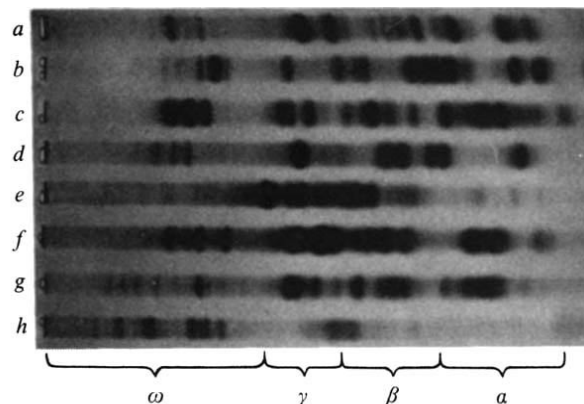


Fig. 1 Polyacrylamide gel electrophoresis¹⁶ of prolamins from different species, aluminium lactate buffer, pH 3.2, migration from left (+) to right (–): a, *T. monococcum*; b, *T. boeoticum* (G-2512); c, *T. urartu* (G-3151); d, *Ae. squarrosa* (0-623); e, *Ae. speltoides*; f, *T. dicoccoides* (64-1148); g, *T. aestivum* ('Scout 66'); h, *S. cereale* ('Frontier'). The α -, β -, γ - and ω -regions correspond to the usual assignments of electrophoretic mobilities for gliadin patterns of common wheats¹.

greater than) the amount of α -type sequence. *T. urartu* G-1876, however, had only a trace amount of Asn at cycle 1; the major amino acid was Val (40 nm) followed by Ala (17 nm). We usually found both Met and Ile at cycle 2, but our sample of *T. monococcum* showed only Ile at cycle 2 whereas our sample of *Aegilops speltoides* showed only Met at cycle 2.

On the basis of N-terminal sequences, our results indicate two major homologous groups of prolamins in the species of *Triticum* and *Aegilops* we examined. Because these species were fairly representative of the two genera, this may be true in general. Based on sequencer yields, we estimate that these two groups account for about 70–80% of our prolamins preparations. A fairly clear difference between the α -type and γ -type sequences occurs at residue 4, where Pro is found in the former and Val or Gln in the latter (Table 2). Pro and Val were the only two amino acids we identified at cycle 4 for all species of *Triticum* and *Aegilops* with the exception of one accession of *T. urartu* (G-1876), which also had a small amount of Leu at this position. The ratio of Pro to Val at cycle 4 may provide a measure of the relative amounts of the two groups of proteins in a sample. These ratios are listed in Table 1, and range from 0.8 to 2.2. They sometimes differed for different varieties or samples of the same species. For example, the two hexaploid bread wheats, 'Scout 66' and 'Justin', had ratios of 1.6 and 0.8, respectively. The large amount of α -type sequence in 'Scout 66' is consistent with the large amounts of α -gliadins in this variety¹¹. In contrast, 'Justin' is low in α -gliadins¹¹. The presence of both α -type and γ -type prolamins in species of *Triticum* and *Aegilops* indicates the close relationship of these genera, and supports the proposal of Morris and Sears¹² that *Aegilops* be incorporated into the genus *Triticum*.

Both the water-soluble and the water-insoluble prolamins from 'Frontier' rye differed from those of the other species examined in lacking the α -type sequence, although we might not have detected a sequence corresponding to about 10% or less of the mixture. The sequences of these rye prolamins preparations were not identical to one another but both were fairly close to the γ -type found in species of *Triticum*⁶ and *Aegilops*. Although minor amino acids were identified at most cycles of the rye prolamins degradation, these did not correspond to known gliadin sequences. The electrophoretic patterns of rye prolamins at pH 3 vary somewhat according to variety¹³, but they do not show a significant amount of staining in the region of mobility corresponding to α -gliadins¹³. As we have noted above, however, the α -type of sequence has been found for proteins with other electrophoretic mobilities⁶, so the agreement between our sequencing results and the electrophoretic patterns may not be essential.

Table 1 Ratio of α -type to γ -type sequence from cycle 4

Species and genome designations	α/γ Ratio
<i>T. aestivum</i> L. em. Thell., 'Scout 66' (ABD)	1.6
<i>T. aestivum</i> L. em. Thell., 'Justin' (ABD)	0.9
<i>T. dicoccoides</i> Körn., 64–1148 (AB)	1.1
<i>T. dicoccum</i> Schubl., Khapli emmer, 63–2356 (AB)	1.3
<i>T. boeoticum</i> Boiss. em. Schiem., G-2512 (A?)	2.2
<i>T. monococcum</i> L., Univ. Manitoba (A?)	1.3
<i>T. urartu</i> Tum., G-3151 (A?, B?)	1.1
<i>T. urartu</i> Tum., G-1876 (A?, B?)	0.7
<i>Ae. speltoides</i> Tausch, U. C., Davis (B?)	1.4
<i>Ae. squarrosa</i> L., 0–623 (D)	1.9
<i>Ae. squarrosa</i> L., 64–1154 (D)	1.9
<i>Secale cereale</i> L., 'Frontier'	—

Table 2 Recovery of PTH amino acids for different species (nmol)

Species	1	2	3	Cycle of Edman degradation					7	8
<i>T. aestivum</i> 'Scout 66'	Val (97) Asn (37)	* Met (40) Ile (22)	Val (65) Gln (54) Leu (12)	Pro (75) Val (47)	Val (69) Asp (28)	Pro (91)	Gln (71) Ser (50)	Leu (64) Gly (37)		
<i>T. aestivum</i> 'Justin'	Val (55) Asn (51)	* Met (42) Ile (40)	Val (50) Gln (39) Leu (32)	Pro (53) Val (63)	Val (41) Asp (24)	Pro (104)	Ser (56) Gln (20)	Leu (46) Gly (42)		
<i>T. dicoccoides</i> 64-1148	Val (64) Asn (33)	* Met (33) Ile (29)	Val (63) Gln (29) Leu (21) Phe (5)	Pro (76) Val (72)	Val (68) Asp (16)	Pro (109)	Ser (70) Gln (25)	Leu (59) Gly (41)		
<i>Ae. squarrosa</i> 64-1154	Val (119) Asn (26) Ala (12)	* Ile (39) Met (17)	Val (91) Gln (25)	Pro (94) Val (47)	Val (53) Asp (9)	Pro (99)	Gln (43)	Leu (39) Gly (10)		
<i>S. cereale</i> 'Frontier' (soluble fraction)	Asn (59) Ala (24) Leu (20)	Met (97) Gln (87)	Gln (95) Leu (78)	Val (35) Asn (15)	Pro (48) Gly (41) Asn (15)	Pro (59) Ser (57)	Ser (73) Gln (14)	Gly (53) Gln (43)		
α -gliadin ^{5,6}	Val	Arg	Val	Pro	Val	Pro	Gln	Leu		
γ_2 -gliadin ⁶	Asn Pro	Ile	Gly Gln	Val	Asp Val Gln	Pro Gln	Trp	Gly Leu		
γ_3 -gliadin ⁶	Asn Pro	Met	Gly Gln	Val Gln	Asp Val	Pro Gln	Trp Gln	Gly		

Recoveries are reported on a basis of 10-mg samples.

* Arg is likely to be an important residue at position 2, but Arg and His were not determined by our GC method: in addition, the quantitation of Asn and Gln was less accurate than that of the other amino acids.

The presence of only the γ -type sequence in our sample of rye prolamins may indicate that rye is closer evolutionarily than *Triticum* and *Aegilops* to the common ancestor that gave rise to the sub-tribe *Triticinae*¹⁴, which includes both wheat and rye, within the family Gramineae. Differentiation may have included the development and amplification of genes coding for α -type prolamins in *Triticum* and *Aegilops*. This speculation is based on results for only one variety of rye; further work on additional varieties and types is needed to allow its generalisation to the species as a whole.

The great homology among the mixtures of prolamins we have studied enables considerable sequencing information to be obtained from them without isolating single protein components. This has also been demonstrated¹⁵ for the zein proteins, the prolamins of maize (*Zea mays*). This homology is notable in consideration of the complexity of the mixtures. The prolamins of species of *Triticum* and *Aegilops* consist of many components that may differ in molecular weight as well as in charge; prolamins preparations from various species of these genera were resolved into at least four (usually more) bands by SDS electrophoresis (E. C. Cole, J. G. Fullington and D.D.K., unpublished results). Small differences in amino acid composition have been found for gliadin components that differ in electrophoretic mobility (aluminium lactate buffer, pH 3.2)¹, and gliadin proteins corresponding to particular bands of the pH 3.2 patterns exhibited simple inheritance in crosses between wheat varieties that differed in their electrophoretic patterns². Regardless of whether the complexity of gliadins and other prolamins mixtures results from multiple genes, post-translational modifications or some other mechanism, the high degree of homology we found for their N-terminal sequences suggests that this region has some important function and thus tends to be conserved.

We thank B. L. Johnson, J. G. Waines, P. Mattern, V. A. Johnson, C. O. Qualset, J. Dvorak, B. L. Jones, and R. Bieter for providing us with grain samples. J.-C.A. is a Visiting Scientist. Reference to a company and/or product name by the Department is only for information and does not imply approval or recommendation to the exclusion of others that may also be suitable.

Received 2 July; accepted 20 September 1979.

1. Kasarda, D. D., Bernardin, J. E. & Nimmo, C. C. in *Advances in Cereal Science and Technology* Vol. 1 (ed. Pomeranz, Y.) 158-236 (American Association of Cereal Chemists, St Paul, Minnesota, 1976).

2. Mecham, D. K., Kasarda, D. D. & Qualset, C. O. *Biochem. Genet.* **16**, 831-853 (1978).
3. Wrigley, C. W. & Shepherd, K. W. *Ann. N. Y. Acad. Sci.* **209**, 154-162 (1973).
4. Hood, L. *Fedn Proc.* **35**, 2158-2167 (1976).
5. Kasarda, D. D., Da Roza, D. A. & Ohms, J. I. *Biochim. biophys. Acta* **351**, 290-294 (1974).
6. Bietz, J. A., Huebner, F. R., Sanderson, J. E. & Wall, J. S. *Cereal Chem.* **54**, 1070-1083 (1977).
7. Charbonnier, L. *C. r. heb. Séanc. Acad. Sci., Paris* **271** Ser. D, 2042-2045 (1970).
8. Edman, P. & Begg, G. *Eur. J. Biochem.* **1**, 80-91 (1967).
9. Pisano, J. J., Bronzert, T. J. & Brewer, H. B. Jr *Analyt. Biochem.* **45**, 43-59 (1972).
10. Niall, H. in *The Proteins*, Vol. III (ed. Neurath, H.) 179-238 (Academic, New York, 1977).
11. Platt, S. G., Kasarda, D. D. & Qualset, C. O. *J. Sci. Fd Agric.* **25**, 1555-1561 (1974).
12. Morris, R. & Sears, E. R. in *Wheat and Wheat Improvement* (eds Quisenberry, K. S. & Reitz, L. P.) 19-87 (American Society of Agronomy, Madison, Wisconsin, 1967).
13. Konarev, V. G., Gavriluk, I. P., Gubareva, N. K. & Peneva, T. I. *Cereal Chem.* **56**, 272-278 (1979).
14. Smith, D. B. & Flavell, R. B. *Biochem. Genet.* **12**, 243-256 (1974).
15. Bietz, J. A., Paulis, J. W. & Wall, J. S. *Cereal Chem.* **56**, 327-332 (1979).
16. Kasarda, D. D., Bernardin, J. E. & Qualset, C. O. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3646-3650 (1976).

Proposed structure of the anthramycin-DNA adduct

Laurence H. Hurley & Ruby Petrussek

College of Pharmacy, University of Kentucky, Lexington, Kentucky 40506

Anthramycin, an antitumour antibiotic produced by *Streptomyces refuineus*¹, belongs to the pyrrolo(1,4)benzodiazepine antibiotic group². Antibiotics within this group are potent inhibitors of nucleic acid synthesis because of their ability to form a labile covalent adduct with DNA²⁻⁶. The reaction of anthramycin with DNA is unusual in that the rate of adduct formation is slow^{2,4,5}, shows an absolute specificity for deoxyguanosine in a double-stranded template⁴, and is only stable as long as the secondary structure of DNA is retained⁶. Experiments using confluent non-dividing human cell lines have demonstrated that while repair-proficient cells are able to remove up to 84% of the adduct within 72 h, xeroderma pigmentosum cells in complementation group A were only able to remove 49% during the same incubation period¹⁰. We propose here a Corey, Pauling and Koltun (CPK) molecular model for the anthramycin-DNA adduct in which anthramycin is attached through the 2-amino group of guanine and lies hidden in the narrow groove. Experimental data supporting this model are provided. The model is based on information about the probable covalent binding sites on anthramycin and guanine together with the known crystal structure of anthramycin⁷.

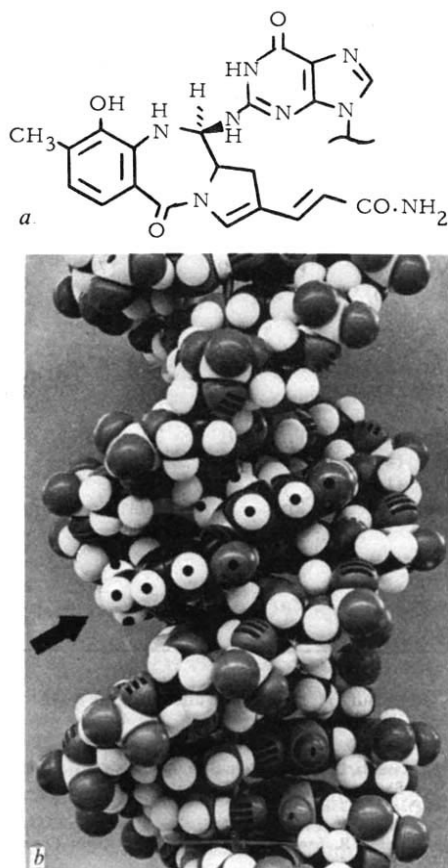


Fig. 1 Structure (a) and CPK space filling model (b) of the anthramycin-DNA adduct in which anthramycin is attached through C-11 to the N-2 group of guanine.

Structure activity relationships for anthramycin and related compounds provide compelling evidence that the carbinolamine is the reactive group on anthramycin^{2,4,5,14}. X-ray diffraction studies on crystals of anthramycin show that the molecule is twisted 40 to 50° from one end to the other along the long axis and thus might fit into one of the grooves of DNA⁷. From studies using synthetic polydeoxynucleotides, guanine is apparently the reactive base on DNA⁴. Of the possible positions for alkylation by anthramycin, N-7, N-3, and C-8 have been eliminated because of: (1) the results of an alkylation assay using [8-³H]guanine-labelled DNA; (2) absence of strand breakage *in vitro* using SV40 DNA; and (3) the lack of increased susceptibility to depurination of the anthramycin modified [8-³H]guanine-labelled DNA⁸. The unreactivity of polydI:dC towards anthramycin, while polydG:dC binds the drug at least as well as native DNA⁴, strongly suggests that the N-2 position on guanine is directly involved in adduct formation. Model building using a CPK molecular model of DNA and anthramycin in the published crystal conformation demonstrated quite dramatically that while O-6 was inaccessible to adduct formation without bond cleavage or serious distortion of DNA, adduct formation through N-2 of guanine produced a remarkably snug fit. The anthramycin molecule which is attached through C-11 to N-2 of guanine by an aminal linkage fits within the narrow groove of DNA (Fig. 1). The predictions which result from this model and the experimental data substantiating our model are presented below.

At saturation binding in DNA, where 25% of the bases are guanine and the drug molecule covers five base pairs, the maximum binding ratio would be 1:10. However, as the distribution of guanine residues is not uniform the frequency will be somewhat lower, and the experimentally determined ratio of 1:12.8 (ref. 2) is probably close to the theoretical value.

Binding of anthramycin should not cause unwinding or lengthening of DNA as the drug binds in a non-intercalative

manner within the narrow groove. Viscosity and sedimentation measurements have shown that the DNA molecule is stiffened but not lengthened on anthramycin binding⁹. Experiments using SV40 DNA, predominantly in the supercoiled form, have shown no unwinding at an antibiotic to base ratio of 1:300.

Our model shows that the angle of the chromophore of anthramycin is inclined at 45° relative to the helix axis. This compared very well with the reported value of 36° as determined by electric dichroism measurement⁹.

The drug is held to DNA by an aminal linkage plus secondary stabilising forces, chiefly between the phenolic group at C-9 of anthramycin and the 2-keto group of cytosine in the same base pair. As the aminal linkage is inherently unstable, conditions which remove the secondary stabilising interaction should lead to release of anthramycin from the adduct. The predictions are mimicked in practice as denaturation of the DNA by acid pH (ref. 5), heat and enzymatic degradation⁶, lead to rapid release of free drug from the DNA adduct.

In the proposed model, anthramycin fits into the narrow groove of DNA without distortion or protrusion outside of the helix. To probe the modified DNA for distortion or protrusion of the adduct outside the helix, S₁ nuclease digestions and benzoylated naphthalated DEAE (BND)-cellulose chromatography on the adducts were carried out. The results (Fig. 2, Table 1) show that S₁ nuclease was unable to reveal any regions of distortion in the drug-modified DNA and, in fact, release of nucleotides was inhibited with increasing amounts of drug modification. The BND-cellulose chromatography results demonstrated lack of single strandedness of DNA induced by anthramycin as well as binding of anthramycin to the resin which would be anticipated if the drug protruded outside the DNA helix.

The orientation of anthramycin in the narrow groove allows interactions between the phenolic group at C-9 of anthramycin and the 2-keto group of cytosine. Titration⁴ and CD measurements⁹ have shown that in the adduct, the phenolic hydroxy proton is no longer available for titration, which is in accord with

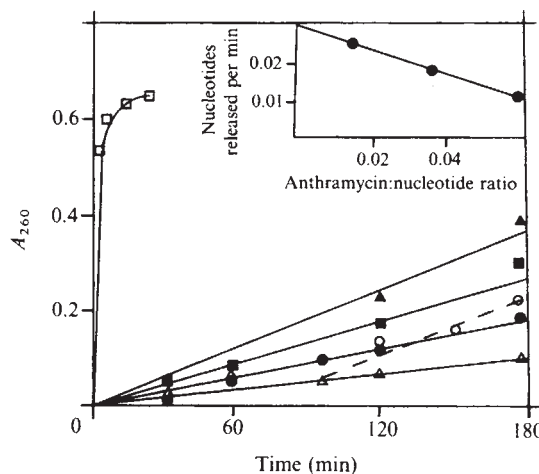


Fig. 2 Time course hydrolysis of calf thymus DNA during digestion with S₁ nuclease in the presence of increasing amounts of bound anthramycin. Assays were carried out essentially as described by Vogt¹³. Calf thymus DNA in 0.1 M SSC was diluted into reaction buffer to produce a final concentration of 0.03M Na acetate pH 4.6, 0.025 M NaCl and 0.005 M ZnSO₄. The reactions were carried out at 37 °C. Aliquots were removed at various times and transferred into 0.4 ml cold 10% perchloric acid. The samples were centrifuged (10,000g for 15 min) and the A₂₆₀ of the supernatant was measured to determine the release of nucleotides by the enzyme. The anthramycin:nucleotide ratio was 0 (▲); 0.015 (■); 0.037 (●); 0.063 (△); DNA concentration 150 µg ml⁻¹, enzyme concentration 2 × 10³ U ml⁻¹. Heat denatured DNA (□). At 90 min unmodified calf thymus DNA was added to an enzyme incubation containing drug-modified calf thymus DNA (○). Inset, dependence of the nucleotides released per min against anthramycin:nucleotide ratio.

Table 1 Estimation of the single strandedness and protrusion of anthramycin molecule external to DNA found in the anthramycin-DNA adduct using BND-cellulose chromatography

Anthramycin: nucleotide ratio	DNA (A ₂₆₀)	% Adherence* ³ H-Anthramycin (c.p.m.)
0	6	—
1:1134	4†	12†
1:144	3	12
1:34	2	10

Chromatography on BND-cellulose in which step elution using 1 M NET (1.0 M NaCl-10⁻⁴ M, EDTA-0.01 M Tris HCl, pH 7.5) eluate and a 50:50 1 M NET: formamide mixture can be used to estimate the double-stranded DNA (eluted with 1 M NET) against DNA containing single-stranded regions¹¹ or protruding aromatic moieties¹² in a sample of DNA. To obtain DNA without single-stranded breaks or regions, calf thymus DNA (Sigma) was pre-chromatographed on BND-cellulose. The 1 M NET eluate (double-stranded DNA) was dialysed against 0.1 SSC, bound to varying concentrations of [15-³H]anthramycin (4.0 µCi µmol⁻¹)⁵, and again dialysed against 0.1 SSC to remove any unbound antibiotic. Control DNA was also dialysed. The DNA-anthramycin adduct (140–220 µg DNA) was then applied to BND-cellulose columns (1 ml column volume) and eluted successively with 10 ml of 0.3 M NET (0.3 M NaCl-10⁻⁴ M, EDTA-0.01 M Tris HCl, pH 7.5), 10 ml of 1.0 M NET and 10 ml of 50% formamide in 1.0 M NET. 2.0 ml fractions were collected. Radioactivity was determined by adding 1.0-ml aliquots to 10 ml of Aquasol (NEN) and DNA by optical density measurement at 260 nm. Optical density measurements of the formamide: 1.0 M NET fractions were corrected for background absorption using a control column without adding DNA.

* The % adherence represents the fraction of absorbance at 260 nm and the fraction of radioactivity eluted with 50% formamide in 1 M NET out of the total amounts eluted with 1 M NET and 50% formamide in 1 M NET.

† The larger values for the apparent percentage adherence found in the radioactivity determinations are most probably due to a small amount of release of drug from the DNA during manipulation before and during chromatography.

our model in which hydrogen bonding between the phenolic group and the 2-keto of cytosine is possible. Our model also accounts for the tighter binding of sibiromycin to DNA than anthramycin as the amino sugar attached at C-7 of sibiromycin binds externally to the deoxyribose phosphate backbone of the helix.

As the anthramycin-DNA adduct is well concealed within the narrow groove of DNA it might be anticipated that excision repair would be a slow process due to the lack of distortion or protrusion which would normally lead to ready recognition of the drug-modified DNA. In practice the drug is only excised relatively slowly over a 48-h period in confluent normal non-dividing human cells¹⁰.

This proposed model for the anthramycin-DNA adduct explains fully both the *in vitro* and *in vivo* data so far obtained. The proposed type of binding to DNA in which the drug molecule is attached by a labile covalent adduct and resides inside the narrow groove seems to be unique to the pyrrolo(1,4)benzodiazepine antibiotics and perhaps explains some of the unusual biological effects of this group of agents. Further experiments using stable isotopically labelled anthramycin and synthetic polydeoxynucleotides should vigorously establish the points of covalent linkage.

We thank Dr Gary Anderson for advice and discussion, Dr Arthur Broom for construction and photographs of the CPK model of the drug adduct, S. Estes, D. Brewer and L. Buch for technical assistance, and the NCI for financial support (CA-17047).

Received 26 June; accepted 26 September 1979.

1. Tendler, M. D. & Korman, S. *Nature* **199**, 501 (1963).
2. Hurley, L. H. *J. Antibiotics* **30**, 349–370 (1977).
3. Kohn, K. W. in *Antibiotics* Vol. III (eds Corcoran, J. W. & Hahn, F. E.) 3–11 (Springer, New York, 1975).
4. Kohn, K. W., Glaubiger, D. & Spears, C. L. *Biochim. biophys. Acta* **361**, 288–302 (1974).

5. Hurley, L. H., Gairola, C. & Zmijewski, M. *Biochim. Biophys. Acta* **475**, 521–535 (1977).
6. Hurley, L. H., Allen, C., Feola, J. & Lubawy, W. C. *Cancer Res.* **39**, 3134–3140 (1979).
7. Mostad, A., Romming, C. & Storm, B. *Acta chim. scand.* **B32**, 639–645 (1978).
8. Hurley, L. H., Buch, L. C., Zimmer, S. G. & Garner, T. F. *Abstr. APhA Acad. pharmac. Sci.* **13**, 8, 112 (1978).
9. Glaubiger, D., Kohn, K. W. & Charney, E. *Biochim. biophys. Acta* **361**, 303–311 (1974).
10. Hurley, L. H., Chandler, C., Garner, T. F., Petrussek, R. & Zimmer, S. G. *J. biol. Chem.* **254**, 605–608 (1979).
11. Scudiero, D., Henderson, E., Norin, A. & Strauss, B. *Mutation Res.* **29**, 473–488 (1975).
12. Karran, P., Higgins, N. P. & Strauss, B. *Biochemistry* **16**, 4483–4490 (1977).
13. Vogt, V. *Eur. J. Biochem.* **33**, 192–200 (1973).
14. Lown, J. W. & Joshua, A. V. *Biochem. Pharmac.* **28**, 2017–2026 (1979).

Energy uptake in the first step of visual excitation

Alan Cooper

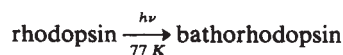
Department of Chemistry, Glasgow University, Glasgow, UK

Perception of light by the retina starts with the absorption of a photon by 11-*cis* retinal, which is covalently incorporated into the membrane-bound protein, rhodopsin¹. The initial result of photon capture is the very rapid formation of a red-shifted species, bathorhodopsin^{2–4} (also known as prelumi-rhodopsin), which is (meta-)stable at liquid nitrogen temperature but which decomposes at higher temperatures, in the dark, through a series of intermediate stages, resulting in the release of all-*trans* retinal from the apoprotein, opsin. Bathorhodopsin formation is the only photochemical step in the overall reaction and, therefore, merits investigation. Several models for the process have been proposed, and have been critically reviewed^{5–8}, although no consensus yet exists as to the nature or mechanism of formation of the batho intermediate. I report here on the first direct measurement of photon energy uptake during bathorhodopsin formation from bovine rhodopsin, and on its possible significance.

We have previously described in detail the technique of photocalorimetry⁹ and its application to the determination of the energetics of the later stages in the photolysis of rhodopsin⁹ and isorhodopsin¹⁰. The present study required the construction of a calorimeter suitable for use at very low temperatures (Fig. 1). This instrument is based on the isothermal heat conduction principle¹¹ and allows direct measurement of the total heat energy dissipated in a sample during monochromatic illumination. In the differential mode, the energy of the exciting light may be subtracted automatically, yielding directly the photochemically-induced uptake or release of energy in the sample. The present instrument has a sensitivity of about 0.1 µV µW⁻¹ at 298 K, falling to about 0.02 µV µW⁻¹ at 77 K due to the intrinsic thermal properties of the thermoelectric sensors. Bovine rod outer segments, prepared by standard sucrose gradient techniques¹² were suspended in either aqueous or neutral detergent (Ammonyx-LO)-glycerol mixtures and irradiated in the photocalorimeter (in liquid nitrogen) at 450 or 485 nm. Samples enriched in bathorhodopsin were generated *in situ* by irradiation at 485 nm and photoreversal was induced by irradiation at 560 nm (ref. 3). Extents of photoreaction were determined by the change in absorbance at 500 nm ($\epsilon = 41,000$) of samples warmed to room temperature and dispersed in 2% (w/v) cetyltrimethylammonium bromide (CTAB), 0.1 M phosphate pH 7.0, 10 mM hydroxylamine. Typically, 10–20 nmol of rhodopsin were bleached in any one experiment.

Illumination of rhodopsin at 77 K gives rise to an unambiguous uptake of energy in parallel with bathorhodopsin formation (Fig. 2a), and this stored energy is released (Fig. 2b) on photoreversal to rhodopsin. (Formation of some 9-*cis* pigment, isorhodopsin, is to be expected during photoreversal³ but, as its energy is only about 5 kcal mol⁻¹ higher than rhodopsin¹⁰ this has no major effect on the observations.) No net heat effects were observed in control experiments with totally bleached rhodopsin samples. The combination of a series of

determinations yields:



$$\Delta H = +34.7 (\pm 2.2) \text{ kcal mol}^{-1}$$

Within experimental error identical results are obtained for irradiation at two different wavelengths (450 and 485 nm), and for both detergent-solubilised rhodopsin and outer segment suspensions. The latter observation suggests that, at this level,

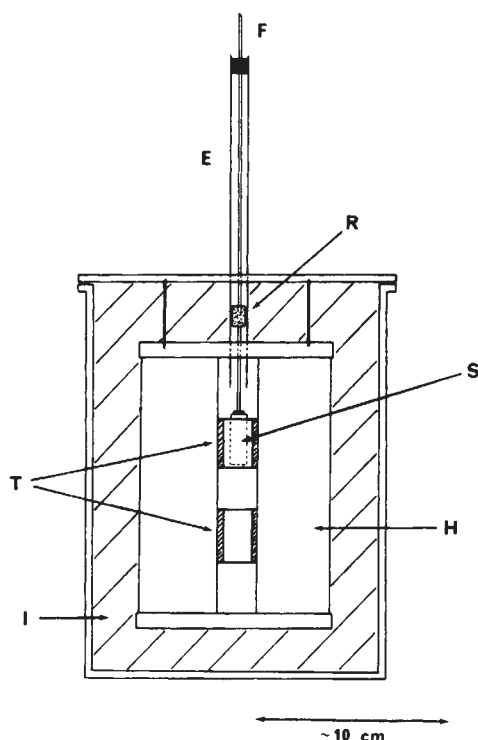


Fig. 1 Low temperature photocalorimeter. The sample (up to 2.5 ml) is enclosed in a cylindrical gold-plated, stainless-steel cell (S) fitted with flexible fibre-optics light guides (F) for illumination. A coaxial thermal radiation shield (R) minimises thermal leakage down the sample entry tube (E). The thermoelectric sensors (T) are solid-state Peltier cooling devices (Borg-Warner 950-71), the output voltage of which is proportional to the rate of heat flow between the sample block and the heat sink (H). The second (lower) set of sensors, with dummy sample block, is connected in opposition to reduce the effects of thermal fluctuations in the heat sink. A second sample cell (not shown) is mounted behind and parallel to the first, such that differential measurements may be made. Small heater coils are permanently mounted in the sample block for electrical calibration. Insulation (I) is with expanded polystyrene foam. The entire unit was suspended in a liquid nitrogen Dewar vessel, with the upper end of E contained in a N₂-flushed dry box to eliminate condensation during sample loading and removal. Light source was a 150 W xenon arc and grating monochromator, with appropriate optics for coupling to the fibre-optics light guides. Sensor voltages were monitored with a Keithley 150B μ V ammeter, averaged and integrated over 1-min intervals, and presented digitally. The equipment was installed in a temperature-controlled darkroom.

the precise molecular environment of the rhodopsin molecule has little influence on the energetics of the primary process.

Thus, the ground-state energy of bathorhodopsin is about 35 kcal (145 kJ, 1.50 eV) higher than rhodopsin. It had been anticipated that the primary reaction should be endothermic, as the thermal instability of bathorhodopsin implies an elevated thermodynamic free energy^{5,9,13}, and entropy contributions are not expected to be significant at low temperatures⁷. The magnitude of the energy storage is, however, much higher than

theoretical models have suggested^{14,15}, amounting to more than 60% of the absorbed photon energy at the 500-nm absorbance maximum of bovine rhodopsin. It is also significantly higher than the measured barrier of ~ 25 kcal for thermal isomerisation about the conjugated double bonds of free retinal¹⁶. If, as seems most probable on evidence so far⁷, bathorhodopsin formation involves at least partial rotation about the retinal 11-12 double bond, then interactions in the retinyl binding site of the protein have a dominating influence on the process, forcing dramatic changes in the conformational potential energy map of the chromophore.

It has been argued that rhodopsin and bathorhodopsin share a common, barrierless, excited-state energy surface⁵⁻⁷. Knowing the ground-state energy of bathorhodopsin, together with spectroscopic and other data, it is possible to begin to reconstruct a more quantitative picture of the potential energy surfaces for the photoreaction (Fig. 3). The interesting possibility then arises that the energy surfaces for the ground and excited states might actually overlap in some region of reaction-space. This conjecture arises from the following considerations: (1) the energy of bathorhodopsin is high (35 kcal above rhodopsin); (2) thermal decay of bathorhodopsin to the next intermediate, lumirhodopsin, involves an activation barrier of about 10 kcal (ref. 17) and, as thermal regeneration of rhodopsin from bathorhodopsin has not been observed, this suggests a ground-state barrier in excess of 45 kcal between rhodopsin and bathorhodopsin; and (3) the excited-state surface must have a minimum below 57 kcal (≈ 500 nm) in order to guide the system towards bathorhodopsin⁵. A hint to the possible intersection or overlap of the energy surfaces comes from the classic experiments of St George¹⁸, which showed that at long wavelengths rhodopsin photolysis becomes temperature dependent, and that

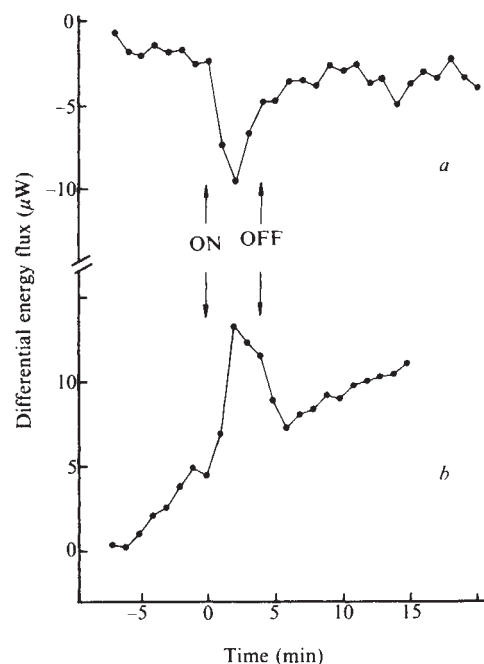


Fig. 2 Differential thermal response on irradiation of rhodopsin and bathorhodopsin in the photocalorimeter at 77 K. The sample consists of bovine rod outer segments dispersed in a mixture of glycerol (67%), Ammonyx-LO (2.5%), 10 mM phosphate buffer pH 7.0. *a*, Rhodopsin $\rightarrow$ bathorhodopsin: 4 min irradiation at 485 nm resulting in a net heat uptake. About 15 nmol of rhodopsin ($\sim 20\%$ of the sample) were bleached in this experiment. *b*, Bathorhodopsin $\rightarrow$ rhodopsin: the above sample was irradiated continuously at 485 nm to produce a steady-state mixture of bathorhodopsin, rhodopsin and isorhodopsin. Subsequent 4 min irradiation at 560 nm (shown) demonstrates the exothermic photoreversal of bathorhodopsin to rhodopsin. (Baseline drift in these experiments is caused by slow temperature variation in the calorimeter, estimated to be less than $5 \times 10^{-6} \text{ }^\circ\text{C min}^{-1}$).

thermal and electromagnetic energy can combine to bleach rhodopsin. The critical wavelength for onset of this phenomenon is 590 nm, suggesting a possible intersection of the energy surfaces at about 48 kcal. Figure 3 has been drawn with this energy in mind.

Given this topology, the response of rhodopsin to photon absorption might be as follows: vertical (Franck-Condon) excitation to a high vibrational level of the excited state followed by rapid vibrational relaxation to lower levels and efficient crossover to isoenergetic states of the ground-state surface in the intersecting region. Further relaxation to the ground-state of bathorhodopsin, or back to rhodopsin, would depend on the dynamics of the situation but, experimentally, bathorhodopsin is favoured in the ratio of about 2:1 (refs 5, 6).

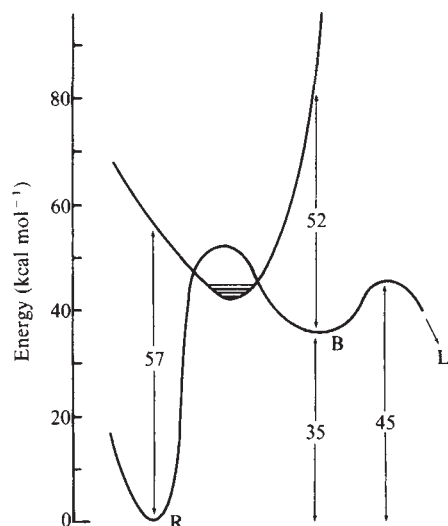


Fig. 3 Two dimensional projection of the possible ground- and excited-state potential energy surfaces for the primary step in rhodopsin photolysis. The ground-state energy of bathorhodopsin (B) with respect to rhodopsin (R) is from direct measurement. Vertical excitation energies to the electronically excited state are from absorbance maxima (500 and 548 nm for R and B, respectively). The B → L (lumirhodopsin) barrier is from kinetic data¹⁷, and suggests a lower limit to the R → B barrier in the ground state. The possible overlap (in projection) of the two surfaces is, probably, exaggerated for clarity. The hatched region indicates possible trapping in lower vibrational levels of the excited-state energy minimum at low temperatures. The nature of the reaction co-ordinate needs further clarification, but probably includes the retinal 11–12 torsion angle^{5–7} as well as other chromophore and protein parameters.

One interesting consequence of energy surface overlap is that at very low temperatures excited molecules might get trapped in vibrational states of lower energy than the ground state of the same conformation (Fig. 3), leading to an energy barrier and a temperature dependence in the rate of bathorhodopsin formation from the excited state, perhaps as observed in picosecond studies¹⁹. However, even at the very lowest temperatures, some crossing over to the ground-state surface might be expected before complete thermal relaxation after excitation, and non-Arrhenius kinetic behaviour might be seen, possibly without the need to invoke quantum mechanical tunnelling processes¹⁹. Deuterium isotope effects¹⁹ would arise in this situation from changes in the vibrational energy levels of modes involving exchangeable hydrogens on the protein and chromophore.

Although we have yet to understand the process in molecular terms, it seems clear that protein–chromophore interactions in rhodopsin have evolved to provide a smooth and efficient pathway for the initial photochemical step in photoreception, whilst maintaining sufficient stored energy to drive the subsequent processes which give rise to the visual stimulus.

I thank Mrs M. Nutley for technical assistance, and various colleagues for helpful discussions. Financial support for this work was provided by the SRC.

Received 16 July; accepted 18 September 1979.

1. Wald, G. *Science* **162**, 230–239 (1968).
2. Yoshizawa, T. & Wald, G. *Nature* **197**, 1279–1286 (1963).
3. Yoshizawa, T. *Handbook of Sensory Physiology* Vol. VII/1, (ed. Dartnall, H. J. A.) 146–179 (Springer, Berlin, 1972).
4. Busch, G. E., Applebury, M. L., Lamola, A. A. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2802–2806 (1972).
5. Rosenfeld, T., Honig, B., Ottolenghi, M., Hurley, J. & Ebre, T. G. *Pure appl. Chem.* **49**, 341–351 (1977).
6. Hurley, J. B., Ebre, T. G., Honig, B. & Ottolenghi, M. *Nature* **270**, 540–542 (1977).
7. Honig, B. A. *Rev. Phys. Chem.* **270**, 540–542 (1977).
8. Honig, B., Ebre, T., Callender, R. H., Dinur, U. & Ottolenghi, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2505–2507 (1979).
9. Cooper, A. & Converse, C. A. *Biochemistry* **15**, 2970–2978 (1976).
10. Cooper, A. *FEBS Lett.* **100**, 382–384 (1979).
11. Wadsö, I. *Acta chim. scand.* **22**, 927–937 (1968).
12. Papermaster, D. S. & Dreyer, W. J. *Biochemistry* **13**, 2438–2444 (1974).
13. Rodieck, R. W. *The Vertebrate Retina* 229–230 (Freeman, San Francisco, 1973).
14. Warshel, A. *Nature* **260**, 679–683 (1976).
15. Warshel, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2558–2562 (1978).
16. Hubbard, R. J. *biol. Chem.* **241**, 1814–1818 (1966).
17. Grellman, K.-H., Livingston, R. & Pratt, D. *Nature* **193**, 1258–1260 (1962).
18. St George, R. C. C. *J. gen. Physiol.* **35**, 495–517 (1952).
19. Peters, K., Applebury, M. L. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3119–3123 (1977).

Evidence for nitrogen fixation (*nif*) genes on indigenous *Rhizobium* plasmids

M. P. Nuti*§, A. A. Lepidi*, R. K. Prakash†, R. A. Schilperoort† & F. C. Cannon‡

* Istituto di Microbiologia Agraria e Tecnica, Università di Pisa, Centro de Studio per la Microbiologia del Suolo, CNR, Pisa, Italy

† Biochemisch Laboratorium, Rijksuniversiteit, Leiden, The Netherlands

‡ ARC Unit of Nitrogen Fixation, University of Sussex, Brighton, UK

The distinguishing characteristic of bacteria that belong to the genus *Rhizobium* is their ability to induce nodules and fix nitrogen in the roots of leguminous plants—hence their importance in agriculture. The demonstration of nitrogen-fixation by rhizobia in the absence of host plant material showed that the genes for nitrogenase synthesis and function reside in the bacterium^{1–3}. The presence of large plasmids (molecular weight (MW) 90–350 × 10⁶) in many *Rhizobium* strains is now well established^{4–6} and earlier genetic evidence suggested that at least some of the genes required for symbiosis were plasmid-linked^{7,8}. More recently, Johnston *et al.*⁹ found evidence that genes required for nodulation were carried on a plasmid in *Rhizobium leguminosarum*. We now report hybridisation between restriction fragments of plasmid DNA from *R. leguminosarum* and DNA encoding part of the structural genes for *Klebsiella pneumoniae* nitrogenase. Our results provide evidence that at least some of the *Rhizobium nif* genes are carried on plasmids indigenous to the strains examined.

Overlapping sequences of *K. pneumoniae nif* DNA which together encode the 15 known *nif* genes have been cloned on small amplifiable plasmids (refs 10 and 11 and unpublished results of M. C. Cannon and F. C. C.). The *nif* genes carried on plasmids pSA30 and pCM1 are shown in Fig. 1. The three structural genes for nitrogenase, *nifKDH* and part of *nifE* are encoded by a DNA fragment (MW 4.2 × 10⁶) cloned in pSA30. An overlapping fragment carrying *nifQ*–*E* and part of *nifK* is

§ Present address: Laboratorio di Microbiologia Agraria, Istituto di Chimica e Industrie Agrarie, Università de Padova, Padua, Italy.

cloned in pCM1 (refs 11, 12). These plasmids were used as radioactive probes to check for the presence of *nif* DNA sequences in *Rhizobium* plasmids. Homology between *nif* DNA from *K. pneumoniae* and *Rhizobium* was expected because: (1) all characterised nitrogenase proteins show a high degree of similarity in their physicochemical properties; (2) component proteins of nitrogenase from different sources form active heterologous enzyme complexes in many of the cases examined; (3) statistical analyses of the amino acid composition of different nitrogenases predicted a high degree of sequence homology¹³.

Plasmid DNA from three strains of *R. leguminosarum* was examined. Strain LPR1705, which is a derivative of the wild-type strain A171, nodulates (*Inf*⁺) *Pisum sativum* effectively (*Eff*⁺)⁵. LPR180, which is *Inf*⁻, was isolated as a non-gummy variant of LPR1705 (ref. 5). JB897 (*phe*, *trp*, *str*), which is a derivative of *R. leguminosarum* strain 300 (ref. 14), nodulates peas effectively. The MWs reported^{5,6} for at least three plasmids detected in strains LPR1705 and JB897 fall within the ranges, 100–110 × 10⁶, 165–200 × 10⁶ and 200–240 × 10⁶. The plasmid of MW 100–110 × 10⁶ was absent in strain LPR180.

The procedure used to investigate *Rhizobium* plasmids for the presence of sequences homologous to *Klebsiella nif* DNA involved: (1) quantitative purification of *Rhizobium* plasmid DNA (refs 15 and 16 and unpublished result of M.P.N. and R.K.P.); (2) digestion of the DNA with restriction endonuclease *Eco*RI; (3) separation of the DNA restriction fragments by agarose gel electrophoresis; (4) denaturation of the DNA fragments in the gel; (5) transfer of the single-stranded DNA bands to nitrocellulose sheets using the Southern technique¹⁷; (6) hybridisation with probes of *Klebsiella nif* or control DNA labelled with ³²P *in vitro*^{16,18}. Figure 2 shows the hybridisation results obtained with the nitrogenase structural gene probe, pSA30. Two DNA bands (MWs 1.67 and 1.33 × 10⁶) from each of strains LPR1705 and LPR180 hybridised to the probe. The hybridisation was weaker with the smaller band from both strains. Three bands (MWs 5.8, 1.67 and 1.4 × 10⁶) from strain JB897 hybridised, one more strongly than the others to pSA30. When pACYC184, the vector for the *nif* genes of pSA30 (Fig. 1), was used as a probe, hybridisation to the above bands was not observed, indicating that the hybridisation of these bands with pSA30 was *nif*-specific. A third band (MW 3.41 × 10⁶) from LPR1705 which hybridised weakly to pSA30 also hybridised to the pACYC184 probe and therefore is probably not *nif* DNA. Further evidence which suggested that the hybridisation was *nif*-specific was obtained from a control experiment in which hybridisation was not observed between pSA30 probe DNA and *Eco*RI digestion fragments of large plasmid DNA from *Agrobacterium tumefaciens* strain Ach5 (ref. 19), a non-nitrogen-fixing member of the Rhizobiaceae. Plasmid pCM1, a

probe which carries part of the nitrogenase structural gene, *nifK* and 11 of the 12 other identified *nif* genes, did not hybridise. This suggests that detectable homology between *Rhizobium* and *Klebsiella nif* DNA is confined to part of the nitrogenase structural genes.

We conclude from these results that the *Rhizobium* plasmid DNA examined contains sequences which are homologous to part of the structural genes for nitrogenase from *K. pneumoniae*. This is strong evidence that at least some of the structural genes for *R. leguminosarum* nitrogenase are plasmid borne. These genes cannot be on the plasmid of MW 100–110 × 10⁶ in LPR1705 because this plasmid is absent in the *Inf*⁻ derivative strain LPR180 which gave an identical hybridisation pattern.

We have also examined plasmid DNA from *Rhizobium meliloti* L5-30 for the presence of sequences homologous to *K. pneumoniae nif* genes. Intensive investigations have revealed only one plasmid of MW 90 × 10⁶ in this strain^{5,6}. Our preliminary results show that two *Eco*RI fragments of DNA (MWs 1.82 and 1.6 × 10⁶) hybridised to the *nif* DNA of pSA30 but not of pCM1. Ruvkun and Ausubel (personal communication) have observed homology between *Rhizobium trifolii* plasmid DNA and the *nif* genes on pSA30, which is

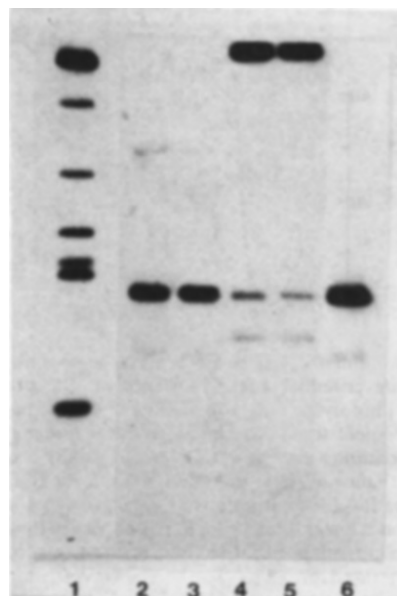


Fig. 2 Autoradiogram of ³²P-labelled pSA30 DNA hybridised to a Southern transfer of control DNA and of *Eco*RI-digested plasmid DNA from *R. leguminosarum* strains. *Eco*RI-generated fragments of *Rhizobium* plasmid DNA, separated by agarose gel electrophoresis were transferred to a nitrocellulose sheet and hybridised to the radioactive probe by a modification¹⁶ of the Southern¹⁷ method. pSA30 probe was labelled *in vitro* to a specific activity of 1–2 × 10⁸ d.p.m. per μg (ref. 16). DNAs on the nitrocellulose sheet were: lane 1, MW standards (5.1, 4.15, 2.65, 2.2, 1.9, 1.69, 0.96 × 10⁶) obtained by combining 100 μg of pSA30 DNA digested with *Eco*RI, 200 μg of pSA30 DNA digested with *Eco*RI + *Hind*III, and 400 μg of pCM1 DNA digested with *Eco*RI; lanes 2–6, plasmid DNA (2 μg) for strains LPR1705 (lanes 2 and 3), JB897 (lanes 4 and 5) and LPR180 (lane 6).

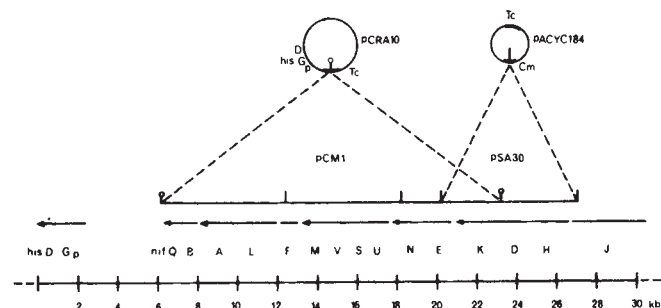


Fig. 1 *Nif* DNA restriction fragments cloned in pCM1 and pSA30 (ref. 9). Symbols for restriction endonuclease sites are: *Hind*III, $\downarrow$ and *Eco*RI, $\perp$. The *his-nif* DNA region and the cloning vectors, pCRA10 and pACYC184, are drawn to the same scale. The arrows indicate *nif* and part of the *his* operons and their direction of transcription—*nifF* and *nifJ* are monocistronic operons. The map of *nif* genes and operon structure are from MacNeil *et al.*²¹, Riedel *et al.*¹² and Merrick *et al.* (in preparation). Tc and Cm are symbols for tetracycline and chloramphenicol resistance genes.

consistent with the transfer of *nif* genes from *R. trifolii* to *K. aerogenes* reported by Dunican and Tierney⁸. These results suggest that the presence of *nif* genes on indigenous plasmids is widespread among *Rhizobium* species. These plasmids can now be investigated for self-transmissibility using the method reported by Beringer *et al.*²⁰ for inserting the transposon Tn5 into *Rhizobium* plasmids.

Our results demonstrate that a combined molecular and genetical analysis of *Rhizobium nif* genes is now possible. Probes of *Klebsiella nif* DNA encoding homologous parts of the

structural genes for nitrogenase could be used to clone *Rhizobium nif* genes, and all the *nif* mutants of *K. pneumoniae* which have been characterised in the past decade are potential hosts for examining such cloned material. These probes could also be used to determine the locations of *nif*:Tn insertion mutations and so to construct a physical map of the *Rhizobium* nitrogenase genes analogous to that of *K. pneumoniae*¹².

We thank Giovanna Bagnoli and Linda Witts for assistance and John Postgate for constructive criticism of the manuscript. This work was supported in part by an EEC grant (no. 441).

Received 12 July; accepted 15 October 1979.

- Pagan, J. D., Child, J. J., Scowcroft, W. R. & Gibson, A. H. *Nature* **256**, 406–407 (1975).
- Kurz, W. G. W. & LaRue, T. A. *Nature* **256**, 407–409 (1975).
- McComb, J. A., Elliott, J. & Dilworth, M. J. *Nature* **256**, 409–410 (1975).
- Nuti, M. P., Ledebor, A. M., Lepidi, A. A. & Schilperoord, R. A. *J. gen. Microbiol.* **100**, 241–248 (1977).
- Prakash, R. K. *et al.* in *Proc. 3rd int. Symp. Nitrogen Fixation*, Madison, 1978 (eds Orme-Johnson, W. H. & Newton, W.) (in the press).
- Casse, F., Boucher, C., Juliot, J. S., Michel, M. & Denarie, J. *J. gen. Microbiol.* **113**, 229–242 (1979).
- Higashi, S. *J. gen. appl. Microbiol.* **13**, 391–403 (1967).
- Duncan, L. K. & Tierney, A. B. *Biochem. biophys. Res. Commun.* **57**, 62–72 (1974).
- Johnston, A. W. B. *et al.* *Nature* **276**, 634–636 (1978).
- Cannon, F. C., Riedel, G. E. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2963–2967 (1977).
- Cannon, F. C., Riedel, G. E. & Ausubel, F. M. *Molec. gen. Genet.* **174**, 59–66 (1979).
- Riedel, G. E., Ausubel, F. M. & Cannon, F. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2866–2870 (1979).
- Eady, R. R. in *The Evolution of Metalloenzymes, Metalloproteins and Related Materials* (ed. Leigh, G. J.) 67–84 (Symposium Press, London, 1977).
- Johnston, A. W. B. & Beringer, J. E. *J. gen. Microbiol.* **87**, 343–350 (1975).
- Ledebor, A. M. thesis, State Univ. Leiden (1978).
- Cannon, F. C. in *Methods for Evaluating Biological Nitrogen Fixation* (ed. Bergersen, F. J.) (Wiley, Chichester, in the press).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
- Klapwijk, P. M., Scheuderman, T. & Schilperoord, R. A. *J. Bact.* **136**, 775–785 (1978).
- Beringer, J. E., Beynon, J. L., Buchanan-Wollaston, A. V. & Johnston, A. W. B. *Nature* **276**, 633–634 (1978).
- MacNeil, T., MacNeil, D., Roberts, G. P., Supiant, M. A. & Brill, W. J. *J. Bact.* **136**, 253–266 (1978).

Inhibition of spinach glyceraldehyde-3-phosphate dehydrogenases by pentalenolactone

Karlheinz Mann & Dieter Mecke

Physiologisch-chemisches Institut, University of Tübingen, Hoppe-Seyler-Strasse 1, D-7400 Tübingen, FRG

The antibiotic pentalenolactone, isolated from the fermentation broth of various species of *Streptomyces*^{1–3}, has three different forms depending on culture conditions and purification procedure: E (epoxide; Fig. 1), D (diol) and C (chlorhydrine). The latter two derive from E by cleavage of the epoxide ring with insertion of the elements of water and HCl, respectively. As *in vitro* studies have shown⁶, pentalenolactone inhibits glycolysis by inhibiting specifically glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), but no other enzyme was known to be affected. The inhibition is irreversible and is similar whether the enzyme comes from *Escherichia coli*, yeast or rabbit muscle. We report here that the antibiotic has different effects on the three classes of glyceraldehyde-3-phosphate dehydrogenases that occur in spinach leaves, and we demonstrate its high specificity for the glycolytic enzyme.

We studied the enzymes of spinach because higher plants contain three different types of dehydrogenases which have glyceraldehyde-3-phosphate as substrate. Two of them are phosphorylating enzymes^{7,8}, distinguishable because one functions in glycolysis, is NAD-specific and located in the cytoplasm (EC 1.2.1.12), while the other is located in the chloroplasts and functions in the dark reactions of photosynthesis, being active with both NAD and NADP (EC 1.2.1.13).

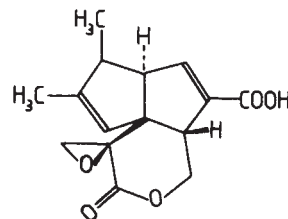


Fig. 1 Structure of pentalenolactone E (refs 4 and 5).

The third catalyses an irreversible and phosphate-independent reduction with NADP as cofactor and is located in the cytoplasm (EC 1.2.1.9; ref. 9).

Pentalenolactone was isolated from the fermentation broth of *Streptomyces arenae* strain Tü 469 by chromatography on XAD-2 resin, silica gel and Sephadex LH-20 as before⁶. Glyceraldehyde-3-phosphate dehydrogenases from spinach were enriched and partially purified, as described by Pupillo *et al.*⁷, by ammonium sulphate fractionation, acetone precipitation, gel filtration, and separation of different enzymes by DEAE-cellulose-chromatography. (Rabbit muscle enzyme and all substrates and cofactors were purchased from Boehringer, Mannheim.)

As Fig. 2 shows, the inhibition of the two phosphorylating enzymes by pentalenolactone was time-dependent. Enzymes were incubated for various times with the antibiotic at room temperature before substrate was added. The activity of both enzymes declined rapidly during the first 2 min, but about 15 min of incubation were necessary for maximal inhibition. We therefore adopted this incubation time in subsequent experiments. Figure 3 shows the concentration dependency of the action of pentalenolactone. Compared with the rabbit muscle enzyme—the most sensitive tested so far—the cytoplasmic NAD-specific spinach enzyme required six times as much antibiotic for 90% inhibition. As Table 1 and Fig. 3 show, this is about the same amount needed to inhibit other glycolytic glyceraldehyde-3-phosphate dehydrogenases. Thus with respect to sensitivity to pentalenolactone, the glycolytic enzymes of the different species belong to the same class. On the other hand, the

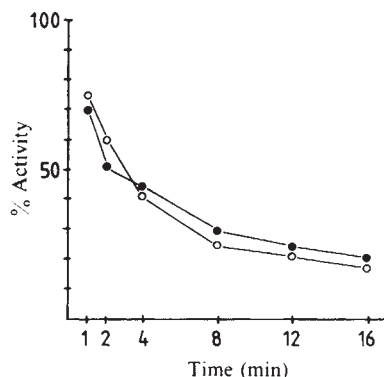


Fig. 2 Time dependency of glyceraldehyde-3-phosphate dehydrogenase inactivation by pentalenolactone. The enzymes were assayed in 0.1 M imidazole buffer, pH 7.4, containing 5 mM EDTA, 5 mM mercaptoethanol, 5 mM sodium arsenate, 2.5 mM coenzyme, and 0.15 IU enzyme in a total volume of 2.1 ml. Pentalenolactone E at a final concentration of 7 μ M (chloroplast enzyme) or 1.5 μ M (cytoplasmic enzyme) was added in 100 μ l of methanol. In a control experiment with 100 μ l pure methanol it was assured, that methanol does not disturb enzyme activity. The assay was started by addition of 250 μ M (final concentration) D-glyceraldehyde-3-phosphate and the absorbancy change at 340 nm was measured for 1 min. The ordinate shows the incubation time before substrate addition in minutes. ○, Cytoplasmic enzyme; ●, chloroplast enzyme.

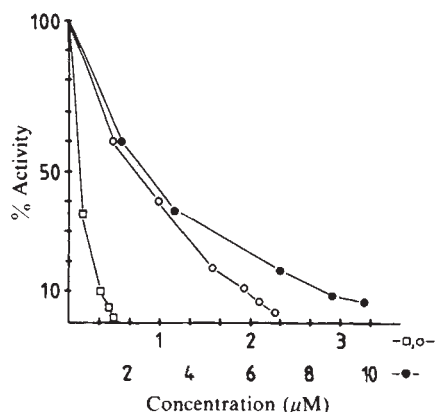


Fig. 3 Concentration dependency of pentalenolactone action. Assay conditions and enzyme concentrations were as detailed in Fig. 2, except for the non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (not shown, see text), when sodium arsenate was omitted. Incubation time before substrate addition was 15 min. □, Rabbit muscle enzyme; ○, spinach cytoplasmic enzyme; ●, spinach chloroplast enzyme. Note the change in scale on abscissa between glycolytic enzymes and chloroplast enzyme.

concentration of antibiotic needed to inhibit the chloroplast enzyme to the same degree is about 25 times those needed to inhibit the rabbit muscle enzyme. Thus, the chloroplast enzyme belongs in another group, as expected from its more variable cofactor dependency. To ensure that the NADP used in the chloroplast enzyme assay was not interfering with inhibition, some assays were done with NAD. The activity of a constant amount of chloroplast enzyme with NAD was about two-thirds of that obtained with NADP, but percentage inhibition with a controlled amount of pentalenolactone was the same. Thus, in a 2.5-μM assay solution the activity of the enzyme with NAD decreased 44% and the activity with NADP 50%. With a 3.5-μM solution 38% of activity was left in both cases. Furthermore, we confirmed that the protein-free plant extract contained nothing that interfered with the action of pentalenolactone. The third glyceraldehyde-3-phosphate converting dehydrogenase of higher plants, the nonphosphorylating enzyme, could not be inhibited to a significant degree by pentalenolactone concentrations up to 150 times that needed to inhibit completely the same amount of rabbit muscle enzyme.

Table 1 Relative sensitivities of different glyceraldehyde-3-phosphate dehydrogenases (GAPDH) to inhibition by pentalenolactone

Enzyme	Relative sensitivity
NAD-GAPDH, rabbit muscle	1*
NAD-GAPDH, <i>E. coli</i>	1.4*
NAD-GAPDH, yeast	3.2*
NAD-GAPDH, spinach	5.7
NAD(P)-GAPDH, spinach	24.5
NADP-GAPDH (nonphosphoryl. spinach)	>150

Sensitivities are expressed as concentration for 90% inhibition divided by concentration for 90% inhibition of muscle enzyme.

* Values from ref. 10.

Our results show that pentalenolactone is a universally applicable inhibitor of glycolysis. The difference in sensitivity to the antibiotic, shown by the three plant enzymes which function in different pathways, stresses the specificity of its action. The binding site of pentalenolactone to glyceraldehyde-3-phosphate dehydrogenase is not known, but Hartmann¹⁰ has shown that the inhibitor strongly interferes with the binding of glyceraldehyde-3-phosphate to the rabbit muscle enzyme. Thus it seems

possible that the substrate-binding part of the enzyme's active site is blocked by the inhibitor. The glycolytic enzymes from various organisms, having the same substrates and cofactors, and consequently a very similar active site, also show the same sensitivity to pentalenolactone. A minor change in the active site, indicated by the ability of the chloroplast enzyme to use NAD as well as NADP, is sufficient to decrease the enzymes' sensitivity to pentalenolactone considerably. Because of the epoxide structure of pentalenolactone E, its irreversible binding shown¹⁰ for the rabbit muscle enzyme and also indicated for the spinach enzymes by the time-dependency of inhibition, and its interference with substrate binding, we conclude that it is covalently bound to the essential thiol group known to react reversibly with glyceraldehyde-3-phosphate during the catalytic cycle. The identification of the pentalenolactone binding site is in progress in our laboratory. We further speculate that the nonphosphorylating plant enzyme, which converts glyceraldehyde-3-phosphate to 3-phosphoglyceric acid, and is completely insensitive to inhibition by pentalenolactone, has no essential sulphhydryl group in its active site.

Received 4 July; accepted 25 September 1979.

1. Koe, K. B., Sobin, B. A. & Celmer, W. D. *Antibiotics Ann.* 1956–1957, 672–675 (1957).
2. Keller-Schierlein, W., Lemke, J., Nyfeler, R. & Zähler, H. *Arch. Mikrobiol.* **84**, 301–316 (1972).
3. Aizawa, S., Akutsu, H., Satomi, T., Kawabata, S. & Sasaki, K. J. *Antibiotics* **31**, 729–731 (1978).
4. Takeuchi, S., Ogawa, Y. & Yonehara, H. *Tetrahedron Lett.* **32**, 2737–2740 (1969).
5. Martin, D. G., Slomp, G., Misak, S., Duchamp, D. J. & Chidester, C. G. *Tetrahedron Lett.* **56**, 4901–4904 (1970).
6. Hartmann, S., Neff, J., Heer, U. & Mecke, D. *FEBS Lett.* **93**, 339–342 (1978).
7. Pupillo, P. & Faggiani, R. *Archs Biochem. Biophys.* **194**, 581–592 (1979).
8. Cerff, R. *Eur. J. Biochem.* **94**, 243–247 (1979).
9. Kelly, G. J. & Gibbs, M. *Pl. Physiol.* **52**, 111–118 (1973).
10. Hartmann, S. thesis, Univ. Tübingen (1979).

Erratum

In the article 'Tholins: organic chemistry of interstellar grains and gas' by C. Sagan and B. N. Khare, *Nature* **277**, 102–107, the following errors, generated at *Nature*, during editing, should be corrected: In paragraph 2 line 2, for H₂C read H₂S. In paragraph 3 line 11, for Sagan⁸ read Sagan¹⁰. In paragraph 4 line 1, experimentally⁹ should read experimentally¹¹. In Table 2 the molecules listed, not the Table itself, are courtesy of Professor L. E. Snyder. The reference for the column labelled 'Gas phase spark' should be 37 not 30, and the reference for the column labelled 'Interstellar frost' should be 11 not 9. The last four lines of Table 2 should read:

Molecule

Propionitrile

6. CH₃CH₂C≡N
6. Cyanodiacyetylene Pentenenitrile-H etc.
CH≡C—C≡C—C≡N
8. Cyanotriacyetylene Benzonitrile ring breakage-H
CH≡C—C≡C—C≡C—C≡N
10. Cyanooctatetrayne
CH≡C—C≡C—C≡C—C≡C—C≡N

In paragraph 6 the penultimate sentence should read: 'The fit of four spectral features between 3 and 4 μm to an accuracy of 0.3 μm will occur by chance with a probability 10⁻⁸'. In paragraph 9 line 13, for tholins³³ read tholins⁴¹. In paragraph 13 line 3, for spluttering read sputtering, and in the same paragraph line 5 should read 'positive ion-molecule reactions may produce the simplest inter-'. In paragraph 15 line 11, for induces³⁴ read induces³⁸. In paragraph 16 line 6, the name of H. C. van de Hulst is misspelled. In reference 5 the authors of the paper in *Astrophys. J. Suppl.* are C. Sagan and E. E. Salpeter.

Corrigendum

In the letter 'Combined effects of adrenaline and insulin on active electrogenic Na⁺-K⁺ transport in rat soleus muscle' by J. A. Flatman and T. Clausen, *Nature* **281**, 580, in Table 3 the units for serum insulin concentration should read μU ml⁻¹.

reviews

Good food for all

Kenneth Mellanby

Food for the Future. By K. Campbell. Pp.178. (University of Nebraska Press: Lincoln and London, 1979.) £8.75.

FOR some years those whom the popular press call ecologists have been telling us that mankind is doomed to perish from starvation. They say that we are not producing enough to feed the present world population of four thousand million, and that the shortage of energy and other resources, and the damage intensive farming is doing to the environment, means that in the future we will produce less, not more food. As the population is expected to be over six thousand million by the year 2000, and as it will probably increase further after that, the future seems desperate. Professor Keith Campbell believes that these doomsday men are "just plain wrong".

He is, of course, not the first to say this. Many agricultural scientists have suggested that, even without fundamental advances in farm practice, vastly greater productivity is possible in many parts of the world. The most striking report, published in 1975 by P. Buringh, H.D.J. van Heemst and G.J. Staring of the Agricultural University, Wageningen in the Netherlands, shows how the world's cereal production might be increased forty-fold. This would feed any possible population increase for many years to come.

Of course everyone, including Professor Campbell, must agree with the Malthusian view that if the world population continues to increase there must eventually be more than can be fed, but there are reasons for believing that this situation is a long way off, or, more hopefully, that population growth will be contained at a satisfactory level (as far as food production is concerned). This view is set out here, using familiar data, based on the fall in birth rate in both developed and developing countries. It is suggested that the most likely estimate is that the world population will stabilise at about eleven thousand million at the end of the twenty-first century; if so there is little doubt that they can all be fed.

Professor Campbell discusses the problems which will affect future agricultural productivity. He shows that

these are of two types. First we have the scientific and agricultural problems — how much food can we produce; will there be land, water, energy and other resources; can the land sustain the present pressure? Secondly we have the economic and political problems — will farmers make a decent living; or will economic and social pressures make them work below their maximum level of productivity?

The first set of problems are dealt with briefly and in a familiar way. Even the most intensive systems use only a tiny fraction of the energy used by man in his other activities, so there should be enough for food production. Global supplies of fertilisers, including phosphate, are adequate for centuries. There is little evidence that food farming, no matter how intensive, is doing anything else than increasing the land's fertility.

Professor Campbell believes that productivity can be greatly increased if research is better organised to deal with the practical problems of the least productive countries. He is critical of much present-day research, particularly in universities, and complains that it is often divorced from the advisory ("extension") services. This view will be endorsed by many older agriculturalists in England and Wales, who regretted the post-War reorganisation which set up the National Agricultural Advisory Service (NAAS), taking advisory work from the universities, cutting off the research workers from the farmers, and producing the usual pullulation of unproductive bureaucrats. Nevertheless I fear that Campbell may be guilty of hubris, for assuming that more relevant research will certainly produce results which will improve food production. Unfortunately we cannot foretell the outcome of original research. But it will probably do some good, even if not as much as is here suggested.

The most interesting sections of this book are those which deal with the other constraints reducing potential food production. The environmentalists who wish to go back to peasant farming are given short shrift; Campbell believes that modern techniques, including chemicals, must be used efficiently, and that there need be no damage to the environment (even if some wildlife may be

exterminated). More serious consideration is given to economic and political policies which encourage low productivity. Price control, quotas and other measures adopted by governments to influence agriculture receive severe and, to me, convincing criticism.

It is clear that poverty, not low food production, is the main cause of hunger in the world today. Wheat prices rose in 1973 because Russia had the money to buy a larger than usual share of the world surplus, thus pricing the poorer countries out of the market. This grain was bought to feed pigs, which had more economic power than starving peasants in the third world. Professor Campbell optimistically suggest that economic growth, world wide, will make increased food production profitable. Some readers will be more sceptical, but they will agree that it will be wise for greater food production, with greater self sufficiency, to be encouraged in poor countries, where internal food movements will be easier than between different sovereign states.

Incidentally, as a good economist, Professor Campbell wishes to give the customer what he wants. He has no time for the suggestion that the amount of animal protein in consumer diets should be reduced. Nevertheless it is realised that where intensively-kept livestock is fed on food which man could consume, more than half the protein and over 90% of the energy is wasted. Today world production is already sufficient to give more than the present population a nutritionally adequate, though dull, largely meatless diet. While we have surpluses, there is no reason why they should not be used to produce luxury food for the richer nations. Whether the rest of the world will ever be able to afford such diets is, in my opinion, doubtful. I am comforted by the knowledge that in a real crisis we could reorganise our feeding habits to ensure that the world's population could be adequately nourished. I only hope that Professor Campbell is justified when he says that "the world can look forward to substantially better diets in the future than were ever enjoyed in the past". □

Kenneth Mellanby was on the staff of Rothamsted Experimental Station from 1955-61, and is author of Can Britain Feed Itself?

Nuclear magnetic resonance

NMR and the Periodic Table. Edited by Robin K. Harris and Brian E. Mann. Pp. 459. (Academic: London and New York, 1978). £35.

THE great bulk of nuclear magnetic resonance work has been done on the nuclei of hydrogen, fluorine, phosphorus, carbon, boron and nitrogen, but most elements of the Periodic Table have at least one nuclide with a magnetic moment. Although quite a lot of work on many elements has been done in a few laboratories in the past twenty years, it is only relatively recently that commercial instruments with good enough resolving power and sensitivity have encouraged their wide use. In this book the editors and an authoritative and experienced group of authors have gathered together most of the work described so far in the literature, together with some unpublished work. The commonly studied nuclei mentioned above are dealt with only briefly and most of the book is devoted to work on the rest of the Periodic Table. The result is a remarkably

comprehensive and fascinating account of an extraordinary range of chemical studies using chemical shifts, nuclear relaxation times, and scalar coupling constants.

The chapters on the various groups of elements also provide tabular data culled from the literature and presented in a convenient way which is likely to be valuable to those working or planning to work on those nuclei; and the coverage of the literature seems to be remarkably comprehensive.

There is a short account of experimental methods, and brief chapters to remind the reader about the theory of the chemical shift, nuclear spin coupling and nuclear relaxation. These are not intended for the beginner nor are they authoritative accounts, but are very useful for reference from time to time when reading other parts of the book.

Dr Harris and Dr Mann are to be congratulated on producing such a useful and interesting book and on having persuaded so many co-authors to cooperate in writing such a compendium.

Rex Richards

Sir Rex Richards, formerly Dr Lee's Professor of Chemistry at the University of Oxford, is now Warden of Merton College and Vice-Chancellor of the University of Oxford, UK.

Geochemical mapping

Regional Geochemical Atlas: Orkney. Compiled by the Institute of Geological Sciences. (Institute of Geological Sciences Bookshop: London, 1979.) £35. *An Introduction to the National Geochemical Data Bank.* (IGS Computer Application Report No. IGS/CAR/78/2.) Second edition. (Institute of Geological Sciences Bookshop: London, 1979.)

THE acquisition of regional geochemical data, based on the collection and analysis of stream sediment samples, had its origin in mineral exploration geochemistry. Individual mining and prospecting organisations would undertake regional geochemical surveys in order to delineate those parts of an exploration area which had the highest probability of containing mineral deposits.

The areas covered by such surveys varied from a few square kilometres to tens or even hundreds of thousands of square kilometres, and great success was achieved with the technique. The data obtained, however, seldom became widely available due to the competitive nature of the mining industry, and as a result some areas in well mineralised environments were surveyed many times. Certain areas in the Irish Republic, for example, are believed to have been surveyed five or six times.

Subsequently, further applications of regional geochemical data have been discovered. The relationship between certain human and animal ailments and the occurrence of abnormally high, or abnormally low levels of certain specific elements in the soils and vegetation of certain areas has been appreciated. The influence of the levels of many elements upon the success or failure of some crops has also been investigated. There can be little doubt that many new uses of regional geochemical data will be discovered in the future, especially if such data become freely available. The publication of the *Regional Geochemical Atlas* by the Institute of Geological Sciences will be especially welcome in this connection.

In the first instance, the atlas will cover the whole of Scotland. The atlas is not the first such publication to appear. As long ago as 1973, the Applied Geochemistry Research Group at Imperial College, London, produced its *Provisional Geochemical Atlas of Northern Ireland*; the somewhat less ambitious *Reconnaissance Base Geochemistry of Parts of the Counties Wicklow, Wexford, Kildare and Dublin* was published by The Geological Survey of Ireland in 1977; and the Applied Geochemistry Research Group published the impressive *Wolfson Geochemical Atlas of England and Wales* (Oxford University Press: Oxford) in 1978.

In the IGS atlas, the distribution of 15 elements in stream sediments is illustrated

in the same number of maps, plus another showing the distribution of uranium in stream waters. Other maps illustrate the relief, geology and mineralisation of the area, together with data collected from a smaller number of sites on the pH and conductivity of the streams themselves.

The maps, on a scale of 1:250000, were produced by automated cartographic methods which were developed jointly by the IGS and the NRTC Experimental Cartography Unit. The level of the relevant element at each sample site is illustrated graphically by the use of lines which are proportional in length to the concentration of the element in the sediment, the lines for each element being orientated in a specific direction.

This method of presentation has the advantage that the results for each sample site are clearly represented, as distinct from the smoothed 'grey-scale' approach adopted by the Imperial College group. Whether or not this manner of data presentation can be considered to be entirely successful, however, is a matter of personal opinion. The reviewer did not find it completely satisfactory, especially where several sample sites occur in close proximity, so that individual lines tend to merge into a single mass which is difficult to interpret.

The prospect of having such regional geochemical maps freely available in the future is clearly one which will increase the scope of geochemical studies. Indeed, one can imagine the time in the future when geochemical maps are regarded as being as basic a requirement as geological maps are today. Nevertheless, many areas will have been the subject of geochemical studies other than those presented in such atlases, and the need to make the data from these studies both available and accessible is also a matter that calls for attention.

The Regional Geochemical Atlas series is part of the Department of Industry's wider Regional Geochemical Reconnaissance Programme under contract to the Natural Environment Research Council. This programme includes the IGS National Geochemical Data Bank which holds data compiled in the production of the geochemical atlases, as well as regional geochemical data from other sources. The establishment of this Data Bank is also to be welcomed, as the use of a central computer for the storage and management of geochemical data could obviously lead to a more effective use of information available.

Data to be accepted initially will be related to igneous, sedimentary and metamorphic rocks, drainage samples (that is, stream sediment samples) and soils, the results for each type of material being stored separately on the basis of geographical location. The bank will be under the control of a manager, who will be responsible for maintaining a suitable level of quality of the data in the bank. Before being archived, data will also be examined by an Advisory Committee consisting of

specialists in each particular field; and during the archiving process, a variety of checks will be applied in order to maintain quality control.

It is perhaps the question of quality control which presents the most important point of contention over the question of a data bank of this type, for not only will the quality of data received vary from project to project, but the level of quality of data that is acceptable to the users of the bank will also vary. Thus a data set that is rejected by the bank as being of unsatisfactory quality, would be denied to a user to whom it might be quite

acceptable. It would perhaps be more satisfactory if data sets were accepted into the bank on the basis of an accompanying quantitative statement of data quality, which would enable the user to select or reject the results of any particular project.

These two publications, therefore, may be seen as heralding a new approach to the study of geochemical problems. It is good to see the IGS leading the way in these important matters.

C. H. James

C. H. James is Senior Lecturer in Economic Geology at the University of Leicester, UK.

Life through Earth's history

The Encyclopaedia of Prehistoric Life. Edited by Rodney Steel and A.P. Harvey. Pp. 218. (Mitchell-Beazley: London, 1979.) £12.

THE editors of an encyclopaedia have a very exacting task in ensuring that their volume lives up to the name, not only in its facts but in its balance and overall content. To ensure factual correctness a formidable array of palaeontologists is here listed as contributors, totalling some 23, most of whom have an international reputation; some of them would no doubt figure on a list of the world's leading palaeontologists, but they hardly constitute such a list, as the dust-cover would have us believe.

The editors are Rodney Steel, a journalist, and Anthony P. Harvey, a librarian; both clearly have an absorbing interest in palaeontology. In their role as editors, however, they have failed to make full use of the talents of their team of contributors, several of whom only have one subject entry, and rely very heavily on their own contributions, which amount to almost 40% of the entries in the book. Rodney Steel contributes on vertebrate palaeontology (in considerable detail), on localities and formations (almost exclusively dealing with vertebrate fossils and omitting, for instance, Ediacara, the Bundenbach Slate, and so on), and on various other subjects with considerably less success. Anthony Harvey uses his bibliographical knowledge to provide biographical information on pioneering palaeontologists, and short sections on institutions with palaeontological museums. Both these lists have some notable omissions but one wonders whether they have place at all in a book on prehistoric life.

There are three main non-editorial contributors, J.W. Franks on plants and L.B. Halstead and R.J.G. Savage on vertebrates. These three provide authoritative reference material which is

amplified by beautifully executed line and stipple drawings which characterise the whole book.

The composition of the rest of the contributing team and the use to which they have been put, raise many queries. One question the logic of using five vertebrate palaeontologists (in addition to Rodney Steel who writes extensively on vertebrates). There are five workers on fossil arthropods, only one of whom has contributed on this phylum (which merits a meagre nine entries); of these five, three work on trilobites but none has contributed to the "trilobite" entry. There are two authors who have worked on the Burgess Shale fauna, but neither has written under that heading. Some 17 of the contributors have produced in total well under 20% of the entries.

The standard of the entries by the professional palaeontologists is high, but it is very clear that some (at least) of the authors have not had the opportunity to check proofs of their own sections. Under "echinoderms", for instance, blastoid hydrospires appear as "hydropores" and "hydrospheres", and an inexcusable editorial insertion of "(spore cases)" appears after the word "theca".

The book is excessively vertebrate-oriented; to widen its appeal, some leaning towards the more conspicuous faunas is excusable. Here though, reasonable limits are exceeded. Over a quarter of the diagrams are of mammals (some groups are dealt with in great detail, so that elephants, mammoths and mastodons have no fewer than eleven large size drawings); another quarter is of reptiles, and a further fifth of other vertebrates. With limited space left for plants and invertebrate groups, the choice of diagrams is sometimes curious. For instance, many echinoderm classes are unfigured, yet there are two large drawings of Lower Jurassic ophiuroids and four views of one species of echinoid.

Molluscs too are poorly served. There is no entry under the heading "cephalopods", although ammonites and belemnites do appear under separate entries. Nautiloids and early ammonoids are described under "ammonites", but coleoid groups such as octopuses,

cuttlefish and squids (the latter two at least with a healthy fossil record) do not merit entries. There are no diagrams of such molluscan groups as scaphopods, polyplacophorans or monoplacophorans, and no mention of segmentation under the latter or under *Neopilina* (mentioned separately, without cross-reference, under "Living fossils").

One of the problems the user of this book must face is inconsistency. Here a firm editorial hand should have spotted these problems. Thus *Stromatoporoids* and *Stromatopora* are treated (by the index at least) as separate organisms; this confusion is compounded by a figured bryozoan being labelled as "*Stromatopora*" rather than *Stomatopora*. Bivalves are referred to as such in the text in some places, and as pelecypods in others — only the latter rates an entry in the index. Ranges of organisms vary too — the first belemnites are stated (correctly) to occur in the Carboniferous under "Belemnites", but in the Permian under the entry for that system. Ranges of some fossil groups are incorrectly stated.

Each geological system has an entry, and most are accompanied by a small time-scale chart showing subdivisions of that system. These are clearly not the authors' work, but are culled from a well-circulated colourful chart which is, unfortunately, full of errors. These are perpetuated in the book, but some new errors have crept in too.

Blackie Biochemical Systematics and Evolution

Andrew Ferguson,
Queen's University, Belfast

This is a review of the most recent developments in biological systematics, written at a level suitable for undergraduates and postgraduates in the biological sciences. The author lays special emphasis on the technique of gel electrophoresis. Throughout, he selects examples which best illustrate particular concepts or applications, although most examples concern animals and particularly the vertebrates.

Suggestions for further reading are provided for each chapter and there are extensive bibliographic citations in the text.

208pp 67 illustrations
Cased ISBN 0 216 90779 9
December 1979 £14.50 net

North American Rights: John Wiley
(Halsted Press)



Blackie & Son Ltd.
Bishopbriggs,
Glasgow G64 2NZ, U.K.

Circle No. 12 on Reader Enquiry Card.

In addition to the numerous line drawings there are fourteen glossy colour charts. Eleven feature vertebrate phylogenies in considerable detail, with many genera mentioned. In contrast two charts of plant and invertebrate phylogenies are very poor and largely at phylum level; the latter lacks a time-scale.

The over-emphasis on vertebrates throughout the book reaches ludicrous proportions when, under the heading "Zone fossils", vertebrate groups receive more space than invertebrates.

There is a glossary which is again heavily vertebrate-oriented. The geological entries are sometimes appallingly incorrect and could not have been written by a professional geologist; the authorship of this section is not given.

The book is presumably aimed at the

layman who wants to learn something about life through Earth's history. If he seeks information on extinct vertebrates, then this may be the book for him, but it will otherwise have a limited appeal. In spite of the excellence of the great majority of the drawings and its wealth of detailed information on vertebrate fossils, its inaccuracies and inconsistencies together with, for most readers, an unacceptable imbalance in its content, render the title *Encyclopaedia* a pretentious one. Though it is attractively laid out and handsomely bound, at £12 for 218 pages it is an expensive way to cover your coffee-table.

J.C.W. Cope

J.C.W. Cope is Senior Lecturer in Geology at University College, Swansea, UK.

Artificial intelligence at MIT

Artificial Intelligence. (Two volumes). Edited by P.H. Winston and R.H. Brown. Pp.492 and 486. (MIT Press: Cambridge, Massachusetts, and London, UK, 1979.) £16.25; \$26 each volume.

AMONG the new scientific arts which have arisen since World War II, artificial intelligence must surely be among the most consequential and the most widely misconstrued. Yet its practitioners are neither diffident nor ambiguous in defining their field of study. "Its central goal", say P.H. Winston and M. Brady in their foreword to a new series from MIT Press, "is to make computers intelligent, both to make them more useful and to understand the principles that make intelligence possible".

The first fruit of the MIT Press series is a two-volume collection of 29 papers by graduate students, young post-doctoral workers and senior staff at the Massachusetts Institute of Technology. They are scattered over a variety of AI-related topics and first appeared, usually in longer form, as technical memos. The editors, P.H. Winston (who now directs MIT's AI Laboratory) and R.H. Brown, have grouped them into sections as follows: in Volume 1 — expert problem solving, natural language understanding and intelligent computer coaches, representation and learning; in Volume 2 — understanding vision, manipulation and productivity technology, computer design and symbol manipulation.

The reader who hopes for a true perspective view, a shining framework of deductively related principles and laws, will recoil from this congeries. Thus some pre-Newtonian seeker after a science of mechanics might have recoiled from

Leonardo's notes on motion, weights, impact and moments of force. Technology's most thrilling advances commonly take place through spasms of untidy thought and activity. While waiting for the orderly abstraction of the theoretician, we may have to quarry general principles from the disorder of notes made in the heat of the struggle by the protagonists themselves. The editors of this book do, however, give conscientious aid to the quarrying process. Very full summarising prefaces precede each section, with an additional shorter note on each chapter. Newcomers to the field will be particularly grateful where specific AI techniques are listed and sign-posted to individual papers.

Some contributions, such as Stallman and Sussman on mechanised circuit analysis, represent deep and thorough finished studies. Others record brave first skirmishes with problems which will not fall before many such assaults have spent themselves. Johan de Kleer's study of qualitative and quantitative reasoning about roller-coaster mechanics is a stimulating example, and Winston's exercise in machine learning through simile raises hope of good eventual solutions in an extremely difficult area. In Marvin Minsky's "The Society Theory of Thinking", in which "the mind is viewed as an organised society of intercommunicating 'agents'", bravery shades into audacity, and the lines of contact with neurobiological fact are stretched exceedingly thin. Yet speculation may still in this field engender aids for the practitioner, as Minsky's earlier "frames" proposal has already amply shown.

Of all the difficult problems addressed by artificial intelligence workers, common consent gives the computational analysis of vision pride of place. Among the six vision papers of the second volume, David Marr's is outstanding. From the strenuous detail and breadth of this attack, methodological lessons can be learned. One of these is

summarised by Winston and Brown: "... it is wrong to begin with devotion to some particular type of algorithm with a view toward finding some problem that it will solve". These words reflect AI's recent change of emphasis for which the MIT school can claim chief credit, namely that intelligent behaviour can only be conjured from a sufficiently rich and massive base of empirical knowledge about the specific domain appropriately organised in machine memory.

Artificial intelligence has thus followed an opposite course to that of the mechanical crafts, which in their early days lost themselves in the properties of specific domains. It was left to Leonardo to make the all-important step of abstracting the concept of a generalised mechanism — for example, for converting between rotary and reciprocating motion — which could then be made to function in any of a variety of different specialised machines. With AI it went the other way, and for 15 years most of us were hypnotised by the promise of general mechanisms — for heuristic search, for deductive reasoning, for machine learning, and the like — as though these universals could accomplish every problem-solving task without the additional need for building in carefully articulated representations of knowledge.

Indeed for a time the MIT school went to the opposite extreme and maintained the exuberant position that domain-specific knowledge should be inextricably interwoven ("procedural embedding") into the substance of the algorithmic mechanisms themselves, instead of maintaining, as we do today, separate "knowledge bases" to which these more general mechanisms can have access. It has now been sufficiently demonstrated that the resulting loss of modularity is fatal to incremental construction and flexible adjustment of an intelligent program's packet of skills. Incremental properties are quintessential. It is refreshing to note from these volumes that the point is now taken, including at MIT.

Space does not allow mention of many other rewarding views to be had of AI people at work. Lozano-Perez's discussion of the problem of telling a robot how to move its hand illuminates the intrinsic complexity of a task at which humans are hereditary Grandmasters. The software technologist will find Hewitt on control as message-passing a source of provoking novelty. Hewitt's chapter closes a book which will irritate purists and antagonise some professionals whose orientation is to Meccas other than MIT. But to those who wish to take a dip into the fast-flowing stream of an active and authoritative laboratory, the publication of this book offers a unique opportunity.

Donald Michie

Donald Michie is Professor of Machine Intelligence at the University of Edinburgh, UK.

obituary

Lew Kowarski

LEW KOWARSKI was the last survivor of the French team that, immediately after the discovery of uranium fission, explored the chances of a nuclear chain reaction. The paper by Otto Hahn and Fritz Strassmann reached Paris about 16 January 1939 and Frédéric Joliot at once saw its tremendous implications. Hans von Halban offered his experience of neutron physics and suggested that Lew Kowarski should join the team. Within weeks they had proof that neutrons, in causing uranium nuclei to split, produced further neutrons; by April they were sure that under suitable conditions those further neutrons would produce further fissions and vice versa, a chain reaction that would lead either to an explosion or to the generation of useful power if the reaction was controlled.

Born in 1907 in St. Petersburg (now Leningrad), Lew had an insecure childhood in a divided household, as son of a Jewish businessman and a Ukrainian singer. At the age of 12 he was smuggled from the Russian revolution to Vilna (in what later became Poland) across the remnants of the German army, some months after his father had gone there. He grew into a giant; musical though he was he had to give up his piano lessons when his fingers grew too thick for the piano keys; so it had to be science. With modest support from his father he studied chemical engineering in Belgium and France.

In 1929 he enrolled in the University of Paris, studying physics and chemistry from books and supporting himself by a job in a firm that made gas pipes, a job he kept part time until 1938. In 1931 he began work on a thesis on crystal growth in Jean Perrin's laboratory, made friends with his son Francis Perrin, and got in touch with Madame Curie, her daughter Irène and Irène's husband Frédéric Joliot just when the couple discovered "artificial" radioactivity at the end of 1933. That brought them the 1935 Nobel prize for chemistry, and Joliot then became professor at the College de France. Kowarski had already been his part-time unpaid assistant for a time; he now became a paid part-time secretary, "Joliot's little typist" as the huge Russian got nicknamed; and then he got a grant that allowed him to drop his other job, of designing gas pipe systems, about which he had just written a book. He still felt an intruder in spite of his doctorate, having got into science by the back door, at the late age of 31; but he was in.

The team agreed that, in publishing, no

ideas should be attributed to individuals; but later it became clear that some of the best ones had been due to Kowarski, who combined inventiveness with meticulous logic. The neutrons were slowed down in water to make them better at causing fission, but many were absorbed before they did. Kowarski urged that heavy water be used, known to absorb hardly any neutrons. It has to be extracted from ordinary water at great cost, and one firm in Norway was doing that; in March 1940 their entire stock, about 40 gallons, was moved to Paris just before the Germans attacked Norway. During that transfer those aliens, Halban and Kowarski, were interned on two separate islands! But time was running out; in May France was invaded, and in June — as instructed by Joliot — Halban and Kowarski escaped to England with their families and the heavy water.

At the Cavendish Laboratory in Cambridge new equipment was built with help from local physicists, mainly under the guidance of Kowarski while Halban negotiated with the USA. The results were promising, but much more heavy water was needed, no longer obtainable from Norway. And why should the USA help Halban, with his German manner, who waved French patents and allied himself with ICI? He seemed more intent to compete after the war than to help win it.

With influential friends and fluent English (though Kowarski soon caught up) Halban assumed a dominant position. In 1942 he was sent to Canada to set up a laboratory; Kowarski, offered an inferior post, decided to stay in Cambridge, as did some of his co-workers. Late in 1943, with some US help promised, (Sir) John Cockcroft took over from Halban, and at last Kowarski got his way: he came to Canada with his team to build a heavy-water reactor. He did it in less than a year; it was the first nuclear reactor outside the USA where Enrico Fermi had achieved the first self-sustained chain reaction in 1942.

After the war Kowarski would have liked to go back to England, where in Cambridge he had felt at home from the first day; but he felt in duty bound to return to France. Under Joliot, scientific head of the Commissariat a l'Energie Atomique, he began to prepare for nuclear power in earnest. Again he built a small heavy-water reactor, both for experimentation and prestige, which went critical late in 1948; nobody knew that the USSR had a reactor going two years before that. Bigger reactors were on the drawing board; would their main product be power or plutonium?

Joliot, now an open communist, publicly refused to make atomic weapons and in 1950 was dismissed from the Commissariat. Kowarski, still working for it part-time, found a new career with CERN, nursing that young European Centre for Nuclear Research. In 1948, his first marriage broken up, he had married Kate Freundlich, daughter of a German scientist; in 1954 they settled in Geneva for good.

As CERN grew his work was gradually taken over by others; in the end he concentrated on data processing with computers, whose importance he saw earlier than most; the device for measuring bubble-chamber tracks that Paul Hough and Brian Powell built with his guidance was used all over the world. A few years' working and travelling for ENEA (the European Nuclear Energy Agency) restored his freedom to visit the USA, denied for years because of his past work with that notorious communist Joliot, and enabled him to hold visiting professorships, two extended ones at Purdue University and shorter ones at Austin (Texas) and at Boston University. In his last years he worked to warn the public against too headlong a growth of nuclear power.

Outwardly like a big jolly Russian peasant, he had a nimble mind and a fine command of both French and English. Honest to the point of bluntness, he made some enemies and many devoted friends. His interest gradually shifted from science to the philosophy and strategy of scientific organisations. Overwork damaged his health, and the single kidney he was born with began to fail years ago; on 27 July he died in a Geneva hospital, cared for by his wife and his daughter Irène from his first marriage. That he survived so long was due to medical art, his own tremendous will, and the care of his calm and gentle Kate.

Otto R. Frisch

Percy Brian

PERCY WRAGG BRIAN died at the age of 68 on 17 August 1979. He was one of the leading botanists of the day with a wide range of interests.

Percy Brian graduated from Cambridge in 1931 with first class honours in the Natural Science Tripos, was awarded the

Frank Smart Studentship as the best student of his year and remained at Cambridge to take a PhD, awarded in 1936. He was awarded an ScD in 1951. His professional career started at Long Ashton Research Station in 1934 and he joined ICI in 1936, where he remained until 1963 — first at Jealott's Hill and later at the Akers Laboratories. He returned to academic life in 1963 as Regius Professor of Botany in Glasgow; in 1968 he moved to Cambridge as Professor of Botany, where he stayed until his retirement in 1977. He was awarded an honorary DSc by Hull University in 1978.

Brian's most productive years as a research worker were spent at the Akers Laboratories. Here he collected an active research group and enjoyed close collaboration with organic chemists who shared his interest in natural products. His research planning and management were faultless; he maintained his interest in a problem only as long as the results justified the work involved; he noticed potential developments peripheral to his central theme, but he did not pursue them until he could give them adequate backing; he did not fritter his energy on red herrings. Once he had launched a new line of research, he gave the workers themselves complete freedom to develop it. The result was a happy, well-integrated and purposeful group.

Brian's central interest as the Akers Laboratories was the influence of antibiotics on the balance of fungi in the soil; this involved collecting antibiotics from soil fungi, examining their effects on other fungi and on higher plants and trying to establish whether the same chemicals were produced in soils. Brian was not deterred by the magnitude and complexity of this problem. During the course of this work he noticed that although most of the antibiotics were generally toxic, a few, notably griseofulvin, were more specifically active against fungi and might have a potential in plant chemotherapy. Brian launched a programme to study this possibility some 10 years before the subject became respectable. Griseofulvin was never successful as a plant chemotherapeutant, but found a unique place in the treatment of mycoses in man and other animals, a possibility which Brian foresaw some years before the chemical was developed.

The chemical interactions of soil fungi included the activity of parasites as well as saprophytes. Among other parasites he studied the products of *Gibberella fujikuroi* and isolated the active chemical, gibberellic acid, in large enough quantities to launch a new area of botanical research. The interactions between fungal parasites and their hosts, particularly as related to the balance of hormones in the plant, remained Brian's central research interest. He pursued these studies at Glasgow and Cambridge with the support of the Agricultural Research Council. The ARC Unit of Developmental Botany was established at Cambridge in

1969 with Brian as honorary director. Here hormone physiology was studied in both healthy and diseased plants.

Brian had less time for research in his university posts but he was known as a tower of strength in his departments and he and his wife, Meg, were untiring in their hospitality to students and to their colleagues.

In addition to research and teaching, Brian was unstinting in the time he would give to the substructure of his science. He was a member of the Agricultural Research Council, twice president of the British Mycological Society, president of the Association of Applied Biologists, the Society of General Microbiology and of the Cambridge Philosophical Society at the time of his death. He was vice-president of the Institute of Biology and served on the council of the Society for Experimental Biology. He also maintained his interest in plant pathology and was one of the chief organisers of the 1st International Congress of Plant Pathology. He was elected a Fellow of the Royal Society in 1958 and served on its Council from 1968 to 1970. He delivered the Leeuwenhoek Lecture in 1966. He was created CBE in 1975.

Brian was never a loquacious man and never pushed himself forward. However, he never refused help and advice when it was requested or needed and what he did say was to the point and well worth listening to. It was probably a measure of his greatness that so many people 'beat a path to his door.' He will be sorely missed by his friends and colleagues and particularly by his family, who must command our deepest sympathy.

S.H. Crowdy

Martin Prestige

MARTIN CALDER PRESTIGE, developmental biologist and Senior Lecturer in Physiology at Edinburgh University, died suddenly on 27 August 1979.

His death, at the early age of 44, removes from the world of developmental neurobiology one of its most incisive minds and able investigators. Educated at Cheltenham College and Trinity College, Cambridge, his basic scientific training was in physiology and led to his being awarded a Michael Foster Studentship. In 1958 he started research, with Patrick Merton, on the occurrence of initial collaterals on kitten motoneurons.

In 1963 Martin joined Arthur Hughes in work on the development of the amphibian spinal cord. This collaboration started in the Anatomy School in Cambridge, continued from 1964 at the Zoology

Department in Bristol, and had a great and continuing effect on the general direction of his work. He took his PhD at Bristol in 1967; this was on the development of the nervous system to the hind leg of *Xenopus* tadpoles and it outlined what was to be one of his main research interests. He continued working in Bristol until he moved to Edinburgh as a Lecturer in Physiology in 1968. He loved Edinburgh and remained there, being appointed Senior Lecturer in 1975.

Two main themes run through Martin's work. The first concerns how the nervous system is put together in terms of cell types and cell numbers; and, with a slight modification of emphasis, this changes into the second theme, which concerns how the system is established in terms of adult functional connectivity.

In Martin's early work his friendship with Arthur Hughes shows most clearly, for he started from Hughes' observations on development in *Xenopus* and extended these brilliantly. In a series of papers starting in 1965 he analysed the phenomena of cell death during the development of the spinal cord and its sensory and motor connections. This was beautiful work, highly relevant to ideas that were then beginning to emerge on the role of central-peripheral relationships in the developing nervous system, and it continued up to his death, with a paper currently submitted for publication.

In 1975 the second main theme of Martin's work was illustrated by the publication, with David Willshaw, of their now well-known paper on theoretical aspects of the formation of retinotectal connections. This paper was very important in that it introduced, for the first time, a welcome touch of rigour into the study of how the eye connects with the brain. Up till then there had been, unbelievably, virtually no serious assessment of the assumptions underlying the various mechanisms that had been proposed to account for the formation of retinotectal connections. Since then there have been several (and more comprehensive) papers on this subject, by others, but the initial credit must go to Prestige and Willshaw.

Martin was also a poet of considerable ability. He had, in his youth, been a mountaineer, and he maintained this interest until his death, in the form of an extensive collection of books on the Himalayas. His first wife, Anne, was killed in a tragic accident. He was married for a second time, to Lucina, in 1972, and their further years in Edinburgh were happy.

In the last year of his life Martin was much troubled by respiratory disability resulting from an illness in youth. Those who knew him over the years will miss him as a friend, as a distinguished scientist, as an inspiration to his colleagues and as an outstanding example of how to behave in the face of adversity.

R.M. Gaze.